# This Page Is Inserted by IFW Operations and is not a part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

## IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.

#### **PCT**

7

#### WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT) (51) Internati nal Patent Classification 5: (11) Internati nal Publication Number: WO 91/00917 C12P 7/62, C12N 15/52 A1 (43) International Publication Date: 24 January 1991 (24.01.91) (21) International Application Number: PCT/US90/03851 (72) Inventors; and (75) Inventors/Applicants (for US only): PEOPLES, Oliver, P. (22) International Filing Date: 10 July 1990 (10.07.90) [US/US]; 104 Fairmont Street, Arlington, MA 02174 (US). SINSKEY, Anthony, J. [US/US]; 285 Commonwealth Avenue, Boston, MA 02115 (US). (30) Priority data: 378,155 10 July 1989 (10.07.89) (74) Agents: PABST, Patrea, L. et al.; Kilpatrick & Cody, 100 Peachtree Street, Suite 3100, Atlanta, GA 30303 (US). (60) Parent Application or Grant (63) Related by Continuation (81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European pa-378,155 (CIP) Filed on tent)\*, DK (European patent), ES (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European 10 July 1989 (10.07.89) (71) Applicant (for all designated States except US): MASSApatent), SE (European patent), US. CHUSETTS INSTITUTE OF TECHNOLOGY (US/ US]; 77 Massachusetts Avenue, Cambridge, MA 02139 (US). Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments (54) Title: METHOD FOR PRODUCING NOVEL POLYESTER BIOPOLYMERS ORF2 encodes a protein of 283 amino acids with an Mr = 31.4 KdPHA depolymerase Transcription EcoR1 EcoR1 pPO23

ORF1 encodes a protein of 559 amino acids with an Mr = 62.3 Kd PHA Polymerase

**ORFI** 

ORF3 encodes a protein of 577 amino acids with an Mr = 64.4 Kd

ORF3

PHA Polymerase

#### (57) Abstract

A method for controlling and modifying biopolymer synthesis by manipulation of the genetics and enzymology of synthesis of polyhydroxybutyrate (PHB) and polyhydroxyalkanoate (PHA) polyesters at the molecular level in procaryotic and eukaryotic cells, especially plants. Examples demonstrate the isolation, characterization, and expression of the genes involved in the production of PHB and PHA polymers. Genes encoding the enzymes in the PHB and PHA synthetic pathway (beta-ketothiolase, acetoacetyl-CoA reductase and PHB polymerase or PHA polymerase) from Zoogloea ramigera strain I-16-M, Alcaligenes eutrophus, Nocardia salmonicolur, and Pseudomonas oleovarans were identified or isolated and expressed in a non-PHB producing organism, E. coli. Specific modifications to the polymers include variation in the chain length of the polymers and incorporation of different monomers into the polymers to produce co-polymers with different physical properties.

ORF2

#### **DESIGNATIONS OF "DE"**

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

#### FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	es	Spain	MC	Monaco
ΝU	Australia	Fí	Finland	MG	Madagascar
BB	Barbados	FR	France	ML	Mali
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Fasso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GR	Greece	NL	Netherlands
BJ	Benin	HU	Hungary	NO	Norway
BR	Brazil	£T.	Italy	PL	Poland
CA	Canada	16	Japan	RO	Romania
CF	Central African Republic	KP	Democratic People's Republic	SD	Sudan
CC	Congo		of Korea	SE	Sweden
CH	Switzerland	<b>€</b> R	Republic of Korea	SN	Senegal
СМ	Cameroon	LI	Liechtenstein	SU	Soviet Union
DЕ	Germany	LK	Sri Lanka	TD	Chad
ÐК	Denmark	LU	Luxembourg	TC	Togo
	•		•	US	United States of America

WO 91/00917 PCT/US90/03851

## METHOD FOR PRODUCING NOVEL POLYESTER BIOPOLYMERS

## Background of the Invention

į

The United States government has rights in this invention by virtue of grants from the National Institute of Health, Office of Naval Research, and National Science Foundation.

5

10

15

20

25

30

Synthesis by bacteria has long been the only means for production of many of the more complex biopolymers. Only recently have pathways for the synthesis of these polymers been determined. Much effort has gone into the isolation of the various enzymes and cofactors involved in these pathways. Regulation of their expression has largely been empirical, i.e., the concentration of nutrients or other factors such as oxygen level have been altered and the effect on polymer production and composition measured.

In order to have control over the production of these complex biopolymers, and to modify them in a specific fashion, it is necessary to design a system for determining the chemical steps required for their synthesis; to isolate and characterize the proteins responsible for these chemical steps; to isolate, sequence, and clone the genes encoding these proteins; and to identify, characterize, and utilize the mechanisms for regulation of the rate and level of the expression of these genes.

Polyhydroxybutyrate, a commercially useful complex biopolymer, is an intracellular reserve material produced by a large number of bacteria. Poly-beta-hydroxybutyrate (PHB), the polymeric ester of D(-)-3-hydroxybutyrate, was first discovered in *Bacillus megaterium* in 1925. Both the chemical and physical properties of this unique polyester have made it an attractive biomaterial for extensive study. PHB has a variety of potential applications, including utility as a biodegradable/thermoplastic material, as a source of chiral centers for the organic synthesis of certain antibiotics, and as a matrix for drug delivery and bone replacement. *In vivo*, the polymer is degraded internally to hydroxybutyrate, a normal constituent of human blood.

10

15

20

25

30

The enzymatic synthesis of the hydrophobic crystalline PHB granules from the C<sub>2</sub> biosynthon acetyl-CoA has been studied in a number of bacteria. Three enzymes: beta ketothiolase, acetoacetyl-CoA reductase and PHB polymerase, are involved in the conversion of acetyl-CoA to PHB.

Thiolases are ubiquitous enzymes which catalyze the synthesis and cleavage of carbon-carbon bonds and thus occupy a central role in cellular metabolism. Different thiolase enzymes are involved in terpenoid, steroid, macrolide and other biosynthetic pathways as well as the degradation of fatty acids. In Z. ramigera, the condensation of two acetyl-CoA groups to form acetoacetyl-CoA is catalyzed by beta-ketothiolase. The acetoacetyl-CoA is then reduced by an NADP-specific reductase to form D(-)-beta-hydroxybutyryl-CoA, the substrate for PHB polymerase.

Beta-Ketothiolase (acetyl-CoA-CoA-C-acetyltransferase, E.C. 2.3.1.9) has been studied in A. beijerinckii (Senior and Dawes, Biochem, J., 134, 225-238 (1973)), A. eutrophus (Oeding and Schlegel, Biochem, J., 134, 239-248 (1973)), Clostridium pasteurianum (Bernt and Schlegel, Arch. Microbiol., 103, 21-30 (1975)), and Z. ramigera (Nishimura et al., Arch. Microbiol., 116, 21-27 (1978)). The cloning and expression of the Z. ramigera acetoacetyl-CoA reductase genes was described in U.S. Serial No. 067,695. This gene was then used as a hybridization probe to isolate the reductase gene from other bacterial species, including Alcaligenes eutrophus and Nocardia.

The reductase involved in PHB biosynthesis in Z. ramigera is stereospecific for the D(-)-isomer of hydroxybutyryl-CoA and uses NADP(H) exclusively as a cofactor. The best characterized Acetoacetyl-CoA reductase is that from Zoogloea, described by Saito et al., Arch. Microbiol., 114, 211-217 (1977) and Tomita et al., Biochemistry of Metabolic Processes, 353, D. Lennon et al., editors (Elsevier, Holland, 1983). This NADP-specific 92,000 molecular weight enzyme has been purified by Fukui, et al., Biochim. Biophys.

25

30

Acta 917, 365-371 (1987) to homogeneity, although only in small quantities. As described in U.S. Serial No. 067,695, the betaketothiolase enzyme from Z. ramigera has now been cloned, expressed and the product thereof purified to homogeneity. The cloned gene was used to identify and isolate the corresponding beta-ketothiolase gene in other bacterial species, including Alcaligenes eutrophus and Nocardia.

The PHB polymerase in Z. ramigera is stereospecific for Dbeta-hydroxybutyryl CoA. Synthetases from other bacteria such as A. eutrophus can utilize other substrates, for example, D-beta-10 hydroxyvaleryl CoA, since addition of propionate into A. eutrophus cultures leads to incorporation of C<sub>s</sub> and C<sub>4</sub> units into a PHB/HV copolymer. Griebel and Merrick, J. Bacteriol., 108, 782-789 (1971) separated the PHB polymerase from native PHB granules of B. megaterium, losing all of the enzyme activity in the process. They 15 were able to reconstitute activity only by adding PHB granules to one of two fractions of the protein. More recently, Fukui et al., Arch. Microbiol., 110, 149-156 (1976) and Tomita et al. (1983), investigated this enzyme in Z. ramigera and partially purified the non-granule bound PHB polymerase. A method for cloning, expressing and using 20 the product thereof in the synthesis of novel polymers was described in U.S. Serial No. 067,695.

A whole range of polyhydroxalkanoate (PHA) storage polymers has been found to be produced by bacteria, including A. eutrophus and P. oleovarans. The PHA polymers are heteropolymers of the D-isomer of β-hydroxyalkanoates with the variation occurring in the length of the side chains (CH<sub>3</sub>-CH<sub>6</sub>H<sub>17</sub>). For example, when grown in the presence of 5-chloropentanoic acid, A. eutrophus incorporates 3-hydroxybutyrate, 3-hydroxyvalerate and 5-hydroxyvalerate into the polymer.

Given the extremely high yields of this polymer obtainable through classic fermentation techniques, and the fact that PHB and

15

20

25

30

PHA of molecular weight greater than 10,000 is useful for multiple applications, it is desirable to develop new PHB-like biopolymers to improve or create new applications.

The production of poly-beta-hydroxyalkanoates, other than PHB, by monocultures of A. eutrophus and Pseudomonas oleovarans was reported by deSmet, et al., in J. Bacteriol., 154, 870-878 (1983). In both bacteria, the polymers were produced by controlled fermentation. A. eutrophus, when grown on glucose and propionate, produces a heteropolymer of PHB - PHV, the PHV content reaching approximately 30%. P. oleovarans produces a homopolymer of polybeta-hydroxyoctanoate when grown on octane. Nocardia has been reported to form copolymers of PHB-PH-2-butenoate when grown on n-butane. Determination of the final composition of 3-hydroxybutyrate polymers by controlled fermentation using selected substrates is also disclosed in U.S. Patent No. 4,477,654 to Holmes et al.

With the availability of a variety of enzymes varying as to their substrate specificity and techniques for expressing the genes encoding the enzymes in other hosts, especially plants, it is possible to provide an economic, biodegradable alternative to the presently available plastics derived from petroleum, especially polypropylene.

It is therefore an object of the present invention to provide further enzymes for use in a method for synthesis of complex biopolymers, particularly PHB, PHA and similar polymers.

It is a further object of this invention to isolate, sequence, and clone additional genes encoding these proteins for polymer synthesis, as well as means for regulation of the rate and level of the expression of these genes.

It is another object of the present invention to provide purified proteins expressed from the genes encoding the proteins for synthesis of polyhydroxybutyrate and polyhydroxyalkanoate.

25

30

It is a further object of the present invention to provide methods for using these proteins and regulatory sequences to create novel biopolymers having polyester backbones.

It is a still further object of the present invention to provide an economic source of biodegradable polyhydroxyalkanoates and novel related polymers, using both bacterial and plant cells for production.

### Summary of the Invention

A method for controlling and modifying biopolymer

synthesis by manipulation of the genetics and enzymology of synthesis
of polyhydroxybutyrate (PHB) and polyhydroxyalkanoate (PHA)
polyesters at the molecular level in procaryotic and eukaryotic cells,
especially plants.

Examples demonstrate the isolation, characterization, and expression of the genes involved in the production of PHB and PHA polymers. Genes encoding the enzymes in the PHB and PHA synthetic pathway (beta-ketothiolase, acetoacetyl-CoA reductase and PHB polymerase or PHA polymerase) from Zoogloea ramigera strain I-16-M, Alcaligenes eutrophus, Nocardia salmonicolor, and Pseudomonas oleovarans were identified or isolated and expressed in a non-PHB producing organism, E. coli.

In a preferred embodiment using bacterial cells, the polymer is made in A. eutrophus due its capacity for accumulating PHB up to 70 to 80% dry cell weigh: under conditions of nitrogen or phosphate limitation. In another embodiment, the genes are introduced into plant cells for expression and synthesis of PHB, PHA, and novel polymers. Specific modifications to the polymers include variation in the chain length of the polymers and incorporation of different monomers into the polymers to produce co-polymers with different physical properties.

10

15

20

25

30

#### Brief Description of the Drawings

Figure 1 is the thiolase gene sequence from Zoogloea ramigera. The sequences located at positions homologous to the E. coli "-10" and "-35" consensus regions (-100 to -95 and -122 and -116) upstream from the transcription start site (bold arrow) are underlined. A probable ribosome binding site is underlined (-11 to -8).

Figure 2 is the complete nucleotide sequence of 2.3 kb of Z. ramigera DNA located downstream from the thiolase gene in clone pUCDBKI, encoding the acetoacetyl CoA reductase. The sequence of 2094 bp extending from the first Sal1 site to the second Sma1 site is shown. Also shown is the translation product of the acetoacetyl-CoA reductase structural gene extending from the ATG at nucleotide 37 to the TGA stop codon at nucleotide 760. Boxed amino acid residues 2 to 6 are identical to those obtained by Edman degradation of the purified protein. A potential ribosome binding site is underlined and a potential transcription terminator is indicated by arrows. Restriction sites for Sal1 and Sma1 are shown.

Figure 3 shows the nucleotide sequence of a corresponding 2 kb fragment A. eutrophus DNA cloned in plasmid pAeT3. The translation products of the A. eutrophus thiolase and acetoacetyl-CoA reductase genes extending from nucleotides 40 to 1219 and 1296 to 2034, respectively, are shown. Restriction endonuclease cleavage sites used in the construction of the overproduction vectors pAT and pAR are shown. Pst 1 = Pst 1; Ava = 2 and Dde = Dde 1.

Figure 4 is the nucleotide sequence of the PHB polymerase (phbC) locus of Alcaligenes eutrophus H16. The translation product of the open reading frame from position 842 to position 2608, the predicted amino acid sequence of PHB polymerase is shown. Also shown is the translation product of the first 72 nucleotides of the phbA gene. A sequence capable of forming a hairpin structure (position 2660) is indicated by the arrows.

10

Figure 5 shows the potential coding regions of the *P. oleovarans* PHA polymerase gene, open reading frames ORF1, ORF2, and ORF3. ORF1 begins at the ATG initiation codon nucleotide 554 and ends at the TGA stop codon nucleotide 2231 and encodes a polypeptide of 559 amino acids with an M<sub>r</sub> = 62,300 daltons. ORF2 begins at the ATG initiation codon nucleotide 2297 and ends at the TAA at 3146, and encodes a 283 amino acid protein with an M<sub>r</sub> = 31,400 daltons. ORF3 begins at the ATG at 3217 and ends at the TGA at 4948, and encodes a protein of 577 amino acids with an M<sub>r</sub> = 64,400 daltons.

Figure 6 is the nucleotide sequence analysis of the complete 6 kb fragment containing the P. oleovarans phbC gene.

## Detailed Description of the Invention

15 The following methods were used to isolate genes encoding beta-keto thiolase, acetoacetyl-CoA reductase, PHB polymerase, and PHA polymerase, and their expression products, to identify and characterize sequences regulating their expression, and to determine the effect of culture conditions and substrate availability on polymer 20 production. Techniques for constructing systems for the production of PHB and PHA-like biopolymers are also disclosed. By combining these enzymes in either bacterial or plant cells with the appropriate substrates under controlled culture conditions of available oxygen and temperature, a variety of polymers can be constructed. The enzymes 25 or nucleotide sequences controlling their expression can also be modified to alter the quantity of expression or substrate specificity to further vary the resulting polymers. An added advantage is that substrates which normally cannot be used with whole cells can be manufactured using the isolated enzymes. 30

The methods, genes, and products of their expression and polymer synthesis are described in detail in the following non-limiting examples.

10

15

20

25

30

#### Media and Culture Conditions.

Zoogloea ramigera strain I-16M (ATCC 19623) was used initially to study the genetics of the PHB biosynthetic pathway. Z. ramigera DNA was purified from 200 ml mid-log phase cultures as follows: cells were harvested by centrifugation, washed in 20 mM Tris-HCl, pH 8.2, and resuspended in 10 ml of Tris-HCl. The cells were then spheroplasted by the addition of 10 ml of 24% w/v polyethylene glycol 8000 and 2 ml of 25 mg/ml lysozyme, followed by incubation at 37°C for 30 min. The spheroplasts were harvested by centrifugation, resuspended in 5 ml of TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA), 300 microliters of 10% w/v SDS added, and the cells lysed by incubating at 55°C for 10 min. An additional 10 ml of TE was added and the lysate incubated with RNAse (50 microgram/ml) and proteinase K (30 microgram/ml) for 1 h at 37°C. The DNA was then purified by CsCl gradient centrifugation.

E coli strains were grown in LB (Luria Bertani) medium, (NaCl, 10 g/l; Tryptone, 10 g/l; yeast extract, 10 g/l) or 2XTY medium (NaCl, 5 g/l; Tryptone, 16 g/l; yeast extract, 10 g/l). For the production of PHB or PHA by E. coli containing recombinant plasmids, minimal media was used, with the modification that the (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> concentration was decreased to 0.04%.

A. eutrophus strains were grown in Trypticase soy broth (TSB, BBL Microbiology systems, Cockeysville, Md) or a defined minimal medium composed of 0.39 g/l MgSO<sub>4</sub>; 0.45 g/l K<sub>2</sub>SO<sub>4</sub>; 12 ml 1.1 m H<sub>3</sub>PO<sub>4</sub>; 15 mg/l FeSO<sub>4</sub>7H<sub>2</sub>O; 24 ml trace elements (20 mg/l CuSO<sub>4</sub>5H<sub>2</sub>O; 100 mg/l ZnSO<sub>4</sub>6H<sub>2</sub>O; 100 mg/l MnSO<sub>4</sub>4H<sub>2</sub>O; 2.6 g/l CaCl<sub>2</sub>2H<sub>2</sub>O). The pH was adjusted to 6.8 with NaOH and the medium sterilized by autoclaving. NH<sub>4</sub>Cl was added to a final concentration of 0.1% or 0.01% as nitrogen source and fructose was added to a final concentration of 0.5-1% (w/v).

Plasmid DNA preparations were carried out using the method of Birnboim and Doly in Nucleic Acids Res., 7, 1513-

ŧ

1523(1979) as described by Ish-Horowicz and Burke, Nucleic Acids Res., 9, 2989-2998 (1981). Lambda DNA was prepared by standard procedures described in Maniatis et al., Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 1982). DNA sequence analysis was carried out using the M13mp18 and M13mp19 cloning vectors (Yanisch-Perron, et al., Gene 33,103-109 (1985)) and the dideoxy chain termination procedure of Sanger, et al., Proc.Natl.Acad.Sci.USA 74, 5463-5467 (1977). α-[35S]-dATP and the Klenow fragment of DNA polymerase 1 were purchased from Amersham. Sequence data were compiled and analyzed on a VAX system.

A recombinant library of random Z. ramigera chromosomal DNA fragments was constructed using the lambda gt11 expression vector described by Young and Davis, Science, 222, 778-782 (1983). Z. ramigera DNA was first methylated using EcoRI methylase and 15 then partially digested with DNAse I in the presence of Mg2, as taught by Anderson, Nucleic Acids, 9, 3015-3026 (1981). The ends of the partially digested DNA fragments were repaired using Klenow polymerase, EcoRI linkers added, and the DNA digested to 20 completion with an excess of EcoRI. Fragments of 2-8 kb were sizeselected on a 1.5% agarose gel, purified by electroelution, and ligated with EcoRI-digested phosphatased lambda gt11 DNA. Ligations were carried out for 18 h at 4°C using 2 micrograms of lambda gt11 DNA and 1 microgram of fragmented chromosomal DNA in a total volume 25 of 10 microliters. The complete ligation reactions were packaged in vitro using lambda extracts prepared from E. coli strains BHB2688 and BHB2690, Hohn and Murray, Proc. Natl. Acad. Sci., USA, 74, 3259-3263 (1977), as described by Maniatis et al. (1982). Packaged phage were plated out and amplified on E. coli Y1088. 30

Screening of the lambda gt11 expression library was carried out using rabbit anti-thiolase antibodies and a modification of the

WO 91/00917 PCT/US90/03851

procedure described by Young and Davis, Science, 222, 778-782 (1983).

Restriction endonucleases, T4 DNA ligase and DNA polymerase 1 were obtained from New England Biolabs and used under conditions provided by the manufacturer. Calf intestinal alkaline phosphatase was purchased from Boehringer Mannheim Corporation. All routine DNA manipulations, including plasmid purifications, E. coli transformations, etc. were performed using methods described by Maniatis, et al. (1982). Chromosomal DNA was purified from A. eutrophus strains, grown to late logarithmic phase in TSB. Transfer of restriction digested DNA samples from agarose gels to nitrocellulose filters, prehybridization and hybridization with <sup>32</sup>P-labelled DNA probes was as described by Peoples, et al., J.Biol.Chem. 262, 97-102 (1987). Rapid plasmid isolation from A. eutrophus recombinant strains for restriction analysis were performed by the alkaline extraction procedure of Birnboim and Doly, Nucleic Acids Res. 7, 1513-1523 (1979).

Conjugation in A. eutrophus.

5

10

15

30

The conjugal transfer of the broad host range plasmid,

pLAFR3, or recombinant derivatives of pLAFR3, into A. eutrophus
was performed using the method described by Easson et al, J.

<u>Bacteriol.</u> 169, 4518-4524 (1987). In this case, however, the recipient
A. eutrophus cells were not sonicated and transconjugants were
selected on A. eutrophus mineral agar plates containing 0.01% NH<sub>4</sub>Cl

as nitrogen source, 1% (w/v) fructose as carbon source and 10 μg/ml
tetracycline.

For Tn5 mutagenesis, a spontaneous streptomycin resistant strain of A. eutrophus 11599 (1599 S1) was used. Transfer of pRK602 (Tn5) was carried out as described above using E. coli MM294A (pRK602) as the only donor. A. eutrophus strains containing Tn5 were selected for by growth on streptomycin (500  $\mu$ g/ml) and kanamycin (100  $\mu$ g/ml).

25

30

## Identification of the Z. ramigera Thiolase gene.

Thiolase antiserum was prepared in New Zealand White female rabbits, using purified thiolase protein by standard procedures. Antibody titer was estimated by the Ouchterlony double-diffusion assay, Acta Pathol. Microbiol. Scand. 26, 507-515(1949). Purified 5 antibody was prepared from the serum by chromatography on protein A agarose according to Bighee et al., Mol. Immunol. 20, 1353-1357 (1983). Approximately 4 x 10 recombinant phage adsorbed to E. coli Y1090 were plated out on 15 cm LB-agar plates and incubated at 42°C for 3 h. The plates were then overlayed with nitrocellulose 10 filters (Schleicher & Schull, BA85), which had previously been saturated in 10 mM IPTG, and incubated a further 4 h at 37°C. Filters were removed, washed for 10 min in TBST (50 mM Tris-HCl, pH 7.9, 150 mM NaCl, 0.05% Tween-20), incubated in TBST plus 15 20% v/v fetal calf serum for 30 min, and rinsed in TBST. First antibody was bound by incubating the filters in 10 ml TBST plus purified anti-thiolase antibody (10 microliters) for 1 h at room temperature. The filters were subsequently washed in three changes of TBST for 5 min each time. Bound first antibody was detected using a biotin-avidin horseradish peroxidase detection system (Clontech Laboratories) and horseradish peroxidase color development reagent (Bio-Rad Laboratories, Richmond, VA).

Proteins were separated by sodium dodecyl sulphatepolyacrylamide gel electrophoresis according to the method of Laemmli, Nature 222, 680-685 (1970) and electrophoretically transferred to nitrocellulose filters (Schleicher & Schuell BA85), essentially as described by Burnette, Anal.Biochem. 112, 195-203 (1981). Following transfer overnight at 30 V, filters were rinsed in TBS (TBST without Tween-20) and incubated in TBS plus 5% bovine serum albumin. Proteins reacting with anti-thiolase serum were then detected by incubating the filters in 100 ml of TBS, 1% gelatin containing 2 ml of anti-thiolase serum for 1-2 h. Bound first antibody

was subsequently detected using goat anti-rabbit IgG horseradish peroxidase conjugate and horseradish peroxidase color development reagent (Bio-Rad Laboratories, Richmond, CA).

DNA blots were prepared using DNA fragments separated on agarose gels by the sandwich blot method of Smith and Summers, 5 Anal. Biochem. 109, 123-129 (1980) based on the technique developed by Southern, J. Mol. Biol., 98, 503-517 (1975). Filters were hybridized with DNA probes labeled to a high specific activity (0.1-1 x 10° cpm/ $\mu$ g of DNA) with [ $\alpha$ -<sup>32</sup>P]dATP, by nick translation, Rigby et al. J. Mol. Biol., 113, 237-251 (1977). Prehybridizations and hybridizations 10 were carried out at 65°C in sealed polythene bags. The prehybridization/hybridization solution contained 5 x SSCP (1 x SSCP contains 0.15 M NaCl, 0.15 M sodium citrate, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 10 mM NaH<sub>2</sub>PO<sub>4</sub>), 5 x Denhardt's solution, 0.1% (w/v) SDS, 10 mM EDTA, and 100 µg/ml sonicated denatured salmon DNA. Filters 15 were prehybridized for 8-18 h and hybridized for 16-18 h using 10' cpm of labeled DNA probe per filter.

Lysogens of lambda gt11 recombinant clones were prepared in E. coli Y1089 as described by Young and Davis, Science 222, 778-782 (1983). For the preparation and analysis of lambda-coded 20 proteins, lysogens were grown at 30°C in LB (100 ml) until they reached an OD of 0.5. The prophage was induced by a 20 min incubation at 45°C, IPTG added to 5 mM and the induced lysogens incubated at 37°C for 1 h. Cells were harvested, resuspended in assay 25 buffer (0.1 M Tris-HCl, pH 7.5, 5 mM beta-mercaptoethanol, 5% (v/v) glycerol), lysed by sonication, cell debris pelleted by centrifugation, and the cell extracts stored at -20°C. The protein concentrations of bacterial lysates were assayed by the method of M. M. Bradford in Anal. Biochem. 72, 248-254 (1976), using bovine serum albumin as a standard. Thiolase-enzyme assays were performed as 30 described by Nishimura, et al., Arch. Microbiol. 116, 21-27 (1978).

10

15

20

25

30

DNA fragments were cloned into the M13 vectors mp10 and mp11 and sequenced by the dideoxy chain-termination method of Sanger et al., Nucleic Acids Res. 10, 141-158 (1980), Proc. Natl. Acad. Sci. USA 74, 5463-5467 (1977). The M13 sequencing primer and other reagents were purchased from Amersham Corp. G/C rich regions were resequenced using dITP in place of dGTP as described by Mills and Kramer, Proc. Natl. Acad. Sci. USA 76, 2232-2235 (1979). Computer-assisted sequence analysis was accomplished using the Staden programs. Nucleic Acids Res. 10, 141-158 (1984).

Approximately 2 x 10<sup>5</sup> recombinants were obtained from 1 μg of purified target DNA, and amplified in *E. coli* Y1088. A total of 10<sup>5</sup> amplified phage were screened using purified rabbit anti-thiolase antibodies. The initial screening identified 10 potentially positive clones (LDBK1-LDBK1O). Analysis by restriction digestions demonstrated that clones LDBK2-10 are identical. Clones LDBK1 and LDBK2 were selected for further study. LDBK1 has an insert composed of 2 EcoRI fragments of 3.6 kb and 0.75 kb. LDBK2 has an insert composed of 3 EcoRI fragments of 1.65 kb and 1.08 kb.

The proteins coded for by the LDBK1 and LDBK2 insert sequences were analyzed both for thiolase-enzyme activity and for cross-reaction to rabbit anti-thiolase serum. Lysogenic strains of *E. coli* Y1089 containing LDBK1 and LDBK2 phage DNA were prepared. Several lysogens were obtained for each clone and two of these, Y1089/LDBK1 and Y1089/LDBK2, were used for subsequent studies. A lysogen of the lambda gt11 vector, BNN97/1ambda gt11, was used as a control. The results of the thiolase-enzyme assays clearly indicate that the proteins from Y1089/LDBK1 contain a substantial amount of thiolase activity. Furthermore, the thiolase activity is inducible, greater than 5-fold, by the addition of IPTG. This shows that expression of the thiolase-coding sequences is under the transcriptional control of the lac promoter contained in the lambda gt11 vector. Neither the Y1089/LDBK2 nor the

30

BNN97/1ambda gt11 protein lysates demonstrate any significant thiolase-enzyme activity even when induced with IPTG.

The size of the proteins responsible for the initial positive reaction to rabbit anti-thiolase antibodies was investigated by Western blot experiments. Protein lysates were separated by SDSpolyacrylamide gel electrophoresis, transferred to nitrocellulose filters, and screened with rabbit anti-thiolase serum. The results show an immunoreactive 40,000 dalton protein in both the IPTG-induced and non-IPTG-induced lysate of Y1089/LDBKI.

10 The LDBKI insert was restriction mapped. The large 3.6 kb EcoRI fragment, containing the complete thiolase gene region, was subcloned into the plasmid vector pUC8 for ease of manipulation. Restriction analysis of one of the subclones obtained, pUCDKB1, confirmed the restriction map of this fragment in LDBKI. pUCDBKI 15 DNA was labeled to a high specific activity with 32P and hybridized to nitrocellulose filters containing Z. ramigera chromosomal DNA digested with the same enzymes used to restriction map pUCDBKI. Southern hybridization experiments confirm that the 5.4 kb genomic fragment hybridizes to both a 1.45 kb SalI/EcoRI and 1.05 kb SalI 20 fragment from pUCDBKl. Based on the result of Southern hybridization experiment, the cloned pUCDBKI insert is represented in the Z. ramigera genome only once.

DNA sequence analysis of the pUCDBKI insert was carried out using the M13/Sanger dideoxy chain termination method. To locate the gene-coding region, individual DNA sequences were 25 scanned in all six reading frames for homology to the NH2-terminal amino acid sequence. By using this approach, the gene-coding region within the 1.45 kb EcoRI/SalI fragment was identified. The complete nucleotide sequence of the plus strand of the gene is shown in Figure 1. 290 bp downstream from the EcoRI site lies the start of the thiolase structural gene, assigned by comparing the DNA sequence to the NH<sub>2</sub>-terminal amino acid sequence. The NH<sub>2</sub>-terminal sequence

10

15

20

25

30

lies in the single long open reading frame which extends from position -89 to the stop codon (TAG) at nucleotide 1174. Beginning with a serine and extending for 25 residues, the experimentally determined NH<sub>2</sub>-terminal sequence aligns identically with residues 2 through 26 of the translated DNA sequence. Translation of the DNA sequence was then used to deduce the remaining amino acid sequence from residue 27 to 391 (nucleotides 79 to 1173). Hence, translation of the DNA sequence from nucleotide 1 to 1174 (in this reading frame) encodes a 391-amino acid polypeptide with a calculated M<sub>r</sub> of 40,598. This value is in very good agreement with that of M<sub>r</sub> = 42,000 determined by SDS-polyacrylamide gel electrophoresis.

Two additional pieces of evidence confirm that this translation produce is the correct amino acid sequence of thiolase. First, a search of the predicted amino acid sequence for the active site peptide (NH<sub>2</sub>-Gly-Met-Asn-Gln-Leu-Cys-Gly-Ser-Gly-COOH) located this peptide at residues 84-92. Finally, the predicted amino acid composition from the translation product and that determined experimentally are in excellent agreement. The G/C content of the 1.46 kb EcoRI-SalI fragment is high, 66.2%. When considered separately, the 5'-flanking 290 bp has a G/C content of 57.4% and the structural gene region 68.4%. The amino acid sequence confirms that the Z. ramigera thiolase contains 5 cysteine residues. The Z. ramigera active site cysteine is at residue Cys-89. Additional cysteines which may be involved in inter- or intradisulphide bonds are Cys-125, Cys-323, Cys-377, and Cys-387. NH<sub>2</sub>-terminal sequence analysis indicated a serine at position 1.

Seven nucleotides upstream from the ATG start codon is a potential ribosome-binding site, 5'-CTGAGGA-3', identified by homology to the *E. coli* sequence. Additional start codons including two GTGs which can initiate translation in some bacterial genes are located further upstream. Examination of the 5'-flanking region for homology to the "-10" and "-35" *E. coli* promoter elements, identified a

the structural gene.

potential "-35' region at residues -122 to -116 and a corresponding "-10 region", 5'-TATAAT-3', at position -100 to -95. A poly(T) tract at position -255 to -266 is also present. It is clear that the only post-translational processing of thiolase is the removal of the N-formylmethionine residue, as alternate start codons, ATC or GTG, are either out of frame or have an in-frame stop codon before the start of

The 1.5 Kb Sall-EcoRl fragment from pUCDBKl contains the entire Z. ramigera thiolase structural gene plus 283 bp of 5'
flanking DNA. A series of plasmid constructions were made in which this DNA fragment was inserted into the tac promoter vector pKK223-3 (or derivatives thereof). pZT3 was made by cleaving pUCKBKl with Sall, blunt-ending with Klenow polymerase and adding on EcoRl linkers. Following digestion with EcoRl, the 1.5 Kb fragment was purified from an agarose gel and inserted into the EcoRl site of pKK223-3. Recombinant clones having the gene inserted in the correct orientation with respect to the tac promoter were identified by restriction analysis following transformation of E. coli JM105.

A series of clones deleted of sequences in the 283 bp 20 flanking the 5' end of the thiolase gene was then constructed. pUCDBKI DNA was digested with EcoRl and treated with Bal31 nuclease. Following Sall digestion, the ends of the fragments were repaired with Klenow and EcoRl linkers added on. The plasmid DNA was cleaved with EcoRl and fragments in the correct size range, 1.2-1.4 kb, purified from an agarose gel and ligated into the EcoRl 25 site of pKK223-3. Clones of interest were identified by restriction mapping and the extent of the 5'-deletion determined by DNA sequencing. From this series of clones, pZT3.1-pZT3.5, the clone with the largest deletion, pZT3.5, had 84 bp of the 5'-flanking DNA 30 remaining and therefore a subsequent Bal31 deletion experiment was carried out as follows: pZT3.5 DNA was digested to completion with EcoRl and treated with Bal31 nuclease; the ends were repaired using

10

15

20

25

30

Klenow polymerase and EcoRI linkers ligated on, following digestion to completion with EcoRI and BamHI, fragments corresponding to the NH<sub>2</sub>-terminal region of thiolase were eluted from an agarose gel, ligated with BamHI - EcoRI digested M13mp 11 DNA and plated out on E. coli JMIOI; single-stranded DNA was prepared from 50 transformants and the extent of the 5'-deletion analyzed by T-tracking; double-stranded DNA was prepared, in vitro, from clones of interest and the EcoRI - BamI inserts recovered by restriction digestion and elution from an agarose gel.

In order to reconstruct the intact thiolase gene, the 290 bp BamHl - HindIII fragment from pZT3.5 was ligated into a vector (pKK226) derived from pKK223-3 by deleting the BamHl site upstream from the tac promoter; this C-terminal vector was subsequently used for the reconstruction of the Bal31 deleted NH<sub>2</sub>-termini of interest; clones pZT3.5.1 and pZT3.5.2 were used in subsequent studies.

The effect of deleting sequences in the 283 bp of DNA flanking the thiolase ATG translation initiation codon was determined by analyzing the level of thiolase activity produced by plasmids pZT3.1-pZT3.5.2. 100 ml cultures of the E. coli JM105 containing each plasmid were induced with IPTG for 15 hours and the level of thiolase assayed. The most notable feature of these results is the very high level of thiolase expression obtained with clones pZT3.3pZT3.5.2, the maximum obtained being 178 U/mg for pZT3.5. This represents an increase of 5.9-fold as compared to plasmid pZT3 which contains the entire 283 bp of 5'-flanking DNA. The data demonstrate that the thiolase 5'-flanking sequences located between -84 (pZT3.5) and -168 (pZT3.2) strongly inhibit the expression of the thiolase gene from the tac promoter. The location of these sequences can be narrowed down to the region between -84 (pZT3.5) and -124 (pZT3.4) as the deletion of this region results in the maximum level of tacdirected thiolase expression. Further deletions to -37 (pZT3.5.1) and

10

15

20

25

30

-26 (pZT3.5.2) do not increase the level of thiolase expression, and in fact a slight decrease is observed. It is important to note that the time course of induction for this series of clones follows the same kinetics as pZT3 and is not appreciably affected by the deletions.

In order to determine if the thiolase promoter lies in the region-84 (pZT3.5) to -124 (pZT3.4), SI nuclease protection experiments were carried out according to the method of Berk and Sharp, Cell 12, 721-732 (1977) on Z. ramigera RNA. Total RNA was isolated from a 100 ml mid-log phase culture by the hot phenol/glass beads extraction procedure of Hinnenbusch et al., J. Biol. Chem. 258, 5238-5247 (1983). 5'-32P-1abelled DNA probe was prepared as follows: 10 µg, plasmid pZT3.1 DNA was digested to completion with Aval and subsequently treated with CIP; the Aval restriction fragments were labelled at the 5'-end with [gamma-32P]-ATP and polynucleotide kinase; following EcoRl digestion, the DNA was separated on an 8% acrylamide gel and the 32P-labelled 280 bp probe fragment eluted and recovered by ethanol precipitation. Probe (10,000 cpm) and 11  $\mu g$ RNA were pooled, freeze dried, resuspended in 10  $\mu$ g hybridization buffer (40 mM pipes, pH 6.4; lmM EDTA, pH 8.0; 0.4 M NaCl; 80% (v/v) formamide), denatured for 10 min at 90°C and annealed at 55°C overnight. 235 microliters ice-cold SI nuclease buffer (0.25 M NaCl; 30mM NaOAc; lmM ZnSO<sub>4</sub>; 200 µg single stranded calf thymus DNA) containing 2000 units of S1-nuclease was added followed by an incubation at 37°C for 30 min. The reaction mixture was extracted once with phenol-chloroform and ethanol precipitated in the presence of 0.2 M NaOAc and 10 µg yeast tRNA carrier. The resulting fragments were analyzed on a 6% (w/v) acrylamide, 7 M urea DNA sequencing gel. For size standards, the Maxam Gilbert G and C sequencing reactions were performed on 50,000 cpm of 5'-32P-labeled probe DNA. The results clearly show a protected fragment and that the RNA start site is located at the C or T residue, position -86 or -87. A control indicates that in the absence of Z. ramigera RNA, the probe is completely degraded, demonstrating the presence of the

10

15

30

thiolase promoter regions approximately 10 bp (-96) and 35 bp (-121) upstream. The 5'-untranslated region of the thiolase is 86 bp long.

From the results of the induction experiments, it is clear that the thiolase gene can be expressed at high levels in a soluble, catalytically active form in E. coli. S1-nuclease studies map the transcription start site for the thiolase gene in Z. ramigera at nucleotides -86/-87.

Studies have demonstrated that although the "-35" region of the thiolase promoter is recognized and binds the RNA polymerase, it is the "-10" region which determines the rate of transcription initiation. In the case of pZT3, for example, the simultaneous binding of an RNA polymerase molecule to both the vector and insert promoters would result in the rapid initiation of transcription from the tac promoter which would subsequently be impeded by the presence of the polymerase molecule bound at the Zoogloea promoter. The closer the two promoters are linked, the less chance of polymerase binding to both at the same time and the lower the inhibition. Therefore, this represents one means for controlling rate of expression of the enzyme. Identification of the Z. ramigera Reductase gene.

20 After identifying the promoter region of the thiolase gene and noting the absence of any potential terminator sequences downstream from the thiolase TAG stop codon, the remaining 2 kb of Zoogloea DNA present in clone pUCDBKI was sequenced and examined for the reductase gene. A series of expression plasmids 25 (pZTI-pZT3) containing either the entire pUDCBKI insert or fragments thereof were constructed in the E. coli tac promoter vector pKK223.3. Each plasmid has the thiolase gene located in the correct orientation for expression from the tac promoter. It is reasonable to expect the tac promoter to direct not only thiolase expression but the expression of any genes located in the 2.3 kb downstream in an operon-like type organization. Clone pZTI was constructed by inserting the entire 3.8 kb EcoRl Z. ramigera DNA insert from

10

15

20

25

30

pUCDBKI into the EcoRl site of the vector pKK223-3. Subsequently, pZT2 was derived from pZTI in a straightforward manner by deleting the 850 bp Smal fragment. pZT3 is constructed as described for the identification of the thiolase promoter. A series of tac promoter induction experiments were performed on each of the recombinant clones pZTI, pZT2 and pZT3. The vector pKK223-3 was used as a control.

E. coli cells containing each of the plasmids were grown and induced by the addition of isopropyl-beta-D-galactopyranoside (IPTG) to a final concentration of 2 mM. After 15 h induction, 10 ml cultures were harvested and lysed by sonication. The cell lysates from each clone were then analyzed both by enzyme assay and on SDS-PAGE. No PHB polymerase activity was detected in any of these lysates. Each of the three recombinant plasmids pZTI, pZT2 and pZT3 demonstrate substantial levels of thiolase activity. In addition, the lysates from pZTI and pZT2 have comparably high levels of AcAc-CoA reductase activity using NADPH as the cofactor. No reductase activity is detected in any of the lysates when NADH is used as a cofactor. The control, pKK223-3, has neither thiolase nor reductase activities. To confirm that the lysates from pZTI and pZT2 do in fact contain the correct reductase, these lysates were also assayed for oxidation of D(-)beta-hydroxybutyryl-CoA. In both cases, enzyme activity was observed with NADP as electron acceptor.

Each of the lysates described above was also analyzed by SDS-PAGE. The results show the presence of the thiolase protein at around 42,000 daltons in protein lysates from pZTI, pZT2 and pZT3, which is not present in the control, pKK223-3. Lysates of pZTI and pZT2 also have a small, 25,000 dalton protein which is not present in the lysate of pZT3 or the control, which corresponds to the AcAc-CoA reductase. The results demonstrate that the AcAc-CoA

reductase. The results demonstrate that the AcAc-CoA reductase gene is located downstream from the thiolase gene. The

entire structural gene for this enzyme must be located between the 3'end of the thiolase and the first Smal site downstream in pUCDBKI.
Identification of the Translation Start Site and Overexpression of the
Reductase Gene.

5

10

The complete nucleotide sequence of the 2339 bp located 2.3 kb downstream from the first Sall site in pUCDBKI is shown in Figure 2. Computer analysis of the sequence data, using codon usage information from the thiolase gene as a standard, identified three open reading frames. N-terminal protein sequence data was obtained from the 25,000 dalton band present in induced lysates from pZTI and pZT2 following preparative SDS-PAGE and electroelution. This data was used to confirm the translation start site for the corresponding gene. The N-terminal five amino acids determined experimentally match residues 2 through 6 predicted from the DNA sequence of the first open reading frame. Translation of this reading frame predicts a polypeptide of 25,000 molecular weight. The translation product of the first open reading frame starting at the ATG, nucleotide 37 and ending at the TGA stop codon nucleotide is shown in Figure 2. This is the predicted primary amino acid sequence of the acetoacetyl-CoA reductase protein.

20

25

30

15

It is evident that the acetoacetyl-CoA reductase gene in clones pZTl and pZT2 can be expressed at reasonably high levels in E. coli. However, in both of these cases, the expression of the reductase gene from the tac promoter is not optimum due to the presence of the thiolase structural gene and 5'-flanking sequence. A simpler acetoacetyl-CoA reductase overproduction vector, pZR14, was constructed. pUCDBKI DNA was digested to completion with Sall and Smal and the Sall ends repaired using the Klenow fragment of DNA polymerase. Following the addition of EcoRl linkers and digestion with EcoRl, the fragments were separated by agarose gel electrophoresis. The 1.05 kb fragment corresponding to the acetoacetyl-CoA reductase structural gene plus 36 bp flanking the 5'-

10

15

25

30

end and 266 bp flanking the 3' end was purified and ligated into the EcoRl site of pKK223-3. pZR14 was then identified as having the correct restriction map with the reductase gene in the right orientation. Induction experiments were performed on pZR14 as described for pZT1, pZT2 and pZT3. Acetoacetyl-CoA reductase was expressed.

Identification of the Thiolase and Reductase Genes in A. eutrophus.

The methods used in isolating the first two PHB genes from Zoogloea were applied to the identification, isolation and characterization of gene from another PHB producing species, Alcaligenes eutrophus, using the Zoogloea thiolase gene region as a hybridization probe to locate homologous sequences.

Subsequent sequence analysis of a 2 kb Pstl fragment of A. eutrophus DNA cloned into pUC8 (clone pAeT3) identified the corresponding thiolase gene region in the A. eutrophus H16 genome. The downstream sequences in pAeT3 are also homologous to the NADP-linked reductase gene region from the Zoogloea clone pUCDBK1. The sequences of the Alcaligenes thiolase and reductase genes is shown in Fig. 3.

Cloning of the individual thiolase and reductase genes from pAeT3 into pKK 223.3, leads to expression of the corresponding enzymes. Comparisons of the Zoogloea and A. eutrophus thiolase protein sequences establish that the two proteins are 63% homologous, including the active site Cys-89.

Both the A. eutrophus and Zoogloea thiolase gene regions were used as hybridization probes to screen Nocardia and Pseudomonas oleovarans DNA for homologous genes. Techniques for identifying the thiolase, reductase, and other synthetase genes from other species having homologous sequences in addition to those described above, are known to those skilled in the art.

Identification of the Z. ramigera PHB polymerase gene.

PHB polymerase from Z. ramigera utilizes D(-)-hydroxybutyryl-CoA monomers, polymerizing them in oxoester linkages in a template-independent head to tail condensation to yield linear polymers. These polymers can contain up to 10,000 monomer units with a molecular weight in excess of 1 x 10<sup>6</sup>. The polymer rapidly becomes insoluble and accumulates as discrete granules in the cell.

A conjugal transfer system based on derivatives of the broad host range plasmid pRK290, described by Ditta et al., in Proc. 10 Natl. Acad. Sci. USA 77, 7347-7351 (1980), transposon mutagenesis and complementation analysis can be used in conjunction with the isolation, characterization and complementation of PHB negative mutants to isolate the PHB polymerase gene for Z. ramigera and A. eutrophus. As described by Schlegel et al., Arch. Microbiol. 71, 283-15 294 (1970), sudan-black staining is used for the detection of PHB negative mutants. Complementation of the mutants is screened for by growing, harvesting, and lysing the cells to release PHB that can then be purified to determine its presence and molecular weight distribution. Thiolase, reductase and PHB polymerase activities in the 20 lysates are also assayed.

Identification of the A. eutrophus PHB polymerase gene.

These techniques were also applied to the cloning, sequencing and expression of the PHB polymerase gene (phbC) in

25 Alcaligenes eutrophus H16 using complementation of poly(B)hydroxybutyrate-negative mutants of A. eutrophus H16. The results demonstrate that the genes encoding the three enzymes of the PHB biosynthetic pathway are organized phbC-phbA-phbB. Expression of all three genes in E. coli results in a significant level (50% dry cell weight) of PHB production in this bacteria. phbC encodes a polypeptide of Mr = 63,900 which has a hydropathy profile distinct

10

15

20

25

3.

from typical membrane proteins indicating that PHB biosynthesis probably does not involve a membrane complex.

The strategy of constructing, characterizing and complementing PHB-negative mutants of a derivative (11599 S1, Table 1) of A. eutrophus H16 was used to identify and isolate the gene(s) encoding PHB polymerase. Transposon mutagenesis allowed use of DNA hybridization analysis to map the chromosomal location of the Tn5 insertion in any interesting strains. 32 potential PHB negative mutants were identified by their opaque colony phenotype when grown on nitrogen limited minimal agar plates. Due to the procedure used to enrich for PHB-deficient strains, it was not surprising that the 32 mutants were found by DNA hybridization using a Tn5 DNA probe to belong to only three classes. More detailed DNA hybridization studies were then used to analyze a representative from each class, i.e., strains PHB #2, PHB #3 and PHB #19. From these studies, it was possible to conclude that in the case of strain PHB #2 and strain PHB #3, the Tn5 insertion causing the opaque phenotype was located in the chromosome approximately 1.2 kb and 1.6 kb, respectively, upstream from the phbA-phbB genes, as illustrated on Figure 3. For strain PHB #19, the Tn5 insertion was located elsewhere on the A. eutrophus chromosome.

The experimental procedure and materials used in the isolation and characterization of <u>phb</u>C were as follows. The procedures and materials are similar to those described for isolation of the <u>phb</u>A and <u>phb</u>B genes.

Bacterial strains and plasmids are shown in Table 1. Media and culture conditions are as described above.

Table 1: Bacterial Strains and Plasmids.

Strain	Relevant Characteristics	Reference
E. coli		Reference
JM83 DH5α	Host strain for plasmids	BRL
A. eutrophus		
H16 11599 11599S1 PHB#2 PHB#3 PHB#19	Wild type strain Strep' H16[phb2::Tn5] H16[phb3::Tn5] H16[phb19::Tn5]	ATCC17699 NCIB 11599
Plasmids .		
pAeT29 pAeT10 pLAFR3 pRK2013 pRK602	phbA-phbB phbA-phbB Tc', cosmid vector Nm' Cm', Nm', pRK2013 nm::Tn9	B. Staskawicz
pUC18 . pUC19	containing Tn5 Ap' Ap'	

DNA manipulations were similar to those described above.

Restriction endonucleases, T4 DNA ligase and DNA polymerase 1
were obtained from New England Biolabs and used as directed by the manufacturer. Calf intestinal alkaline phosphatase was purchased from Boehringer Mannheim Corporation. All routine DNA
manipulations, including plasmid purifications, E. coli transformations, etc. were performed using methods described by Maniatis, et al.
Chromosomal DNA was purified from A. eutrophus strains, grown to late logarithmic phase in TSB as described previously. Transfer of restriction digested DNA samples from agarose gels to nitrocellulose filters, prehybridization and hybridization with <sup>12</sup>P-labelled DNA probes was as previously described. Rapid plasmid isolation from A.

10

15

eutrophus recombinant strains, for restriction analysis, was performed by the alkaline extraction procedure.

## Conjugation in A. eutrophus.

The conjugal transfer of the broad host range plasmid, pLAFR3, or recombinant derivatives of pLAFR3, into A. eutrophus was performed using the same method as previously described. In this case, however, the recipient A. eutrophus cells were not sonicated and transconjugants were selected on A. eutrophus mineral agar plates containing 0.01% NH<sub>4</sub>Cl as nitrogen source, 1% (w/v) fructose as carbon source and 10  $\mu$ g/ml tetracycline.

For Tn5 mutagenesis, a spontaneous streptomycin resistant strain of A. eutrophus 11599 (1599 S1) was used. Transfer of pRK602 (Tn5) was carried out as described above using E. coli MM294A (pRK602) as the only donor. A. eutrophus strains containing Tn5 were selected for by growth on streptomycin (500  $\mu$ g/ml) and kanamycin (100  $\mu$ g/ml).

## Amplification and Identification of PHB-Deficient Mutants.

The amplification and screening procedures described by Schlegel and Oeding, Radiation and Radioisotopes for Industrial 20 Microorganisms. International Atomic Energy Agency, Vienna, 223-231 (1971) was used to identify PHB-deficient strains of A. eutrophus. A pool of around 10<sup>s</sup> Kan' transconjugants (Tn5 insertion mutants) was inoculated into 10 ml of mineral media containing 0.01% NH<sub>4</sub>Cl, 1% fructose and 100  $\mu$ g/ml kanamycin and incubated for 18 h at 25 30°C. This culture was then used to inoculate 100 ml of the same medium and incubated for 30 h at 30°C. To amplify PHB-deficient mutants, aliquots of this culture containing approximately 10° cells were fractionated on sucrose step gradients by density equilibrium centrifugation and plated out on mineral agar plates containing 0.01% NH<sub>4</sub>Cl, 1% fructose and 100  $\mu$ g/ml kanamycin. After growth for 4-5 30 days at 30°C, opaque (PHB-deficient) and white (PHB-containing) colonies were readily distinguished. By quantitating the level of PHB produced by both opaque and white colonies, it was confirmed that

opaque colonies were PHB-deficient whereas white colonies contained PHB.

#### Analysis of Proteins.

In order to perform assays for B-ketothiolase, NADPHlinked acetoacetyl-CoA reductase and PHB-polymerase, 100 ml 5 cultures of A. eutrophus strains were grown at 30°C for 40 hours in TSB. For Tn5 mutant strains, kanamycin was added at 100  $\mu$ g/ml and for strains containing pLAFR3 or derivatives thereof, tetracycline was added to 10 µg/ml. Cells were harvested by centrifugation, resuspended in 2 ml lysis buffer (10 mM Tris HCl, pH 8.0; 5 mM B-10 mercaptoethanol; 5 mM EDTA, 0.02 mM phenyl-methyl-sulfonylfluoride; 10% v/v glycerol) and lysed by sonication. An aliquot of the lysate was cleared of cell debris by centrifugation for B-ketothiolase and acetoacetyl-CoA reductase assays. B-ketothiolase activity was 15 determined by measuring the rate of thiolysis of acetoacetyl-CoA as described by Davis, et al., J. Biol.Chem. 262,82-89 (1987), of acetoacetyl-CoA with NADPH as the cofactor. PHB polymerase assays were performed using samples of the crude lysate and determining the level of incorporation of D-3H-hydroxybutyryl-CoA (specific activity approximately 2  $\mu$ Ci/ $\mu$ mol), as described by Fukui, et 20 al., Arch. Microbiol. 110, 149-156 (1976). Protein concentrations were determined by the method of Bradford, Anal. Biochem. 72, 248-254 (1976), using Biorad assay solution and bovine serum albumin as the standard. E. coli maxi-cell labelling studies were performed as described by Sancar, et al., J. Bacteriol. 137, 692-693 (1979). 25

## PHB Purification and Quantitation.

To determine the level of PHB in different strains, 100 µl aliquots of the crude lysates were treated with 1.2 ml of 5% Na hypochlorite solution for 1 h at 37°C. The insoluble PHB was then harvested by centrifugation for 10 min in a microcentrifuge, washed successively with 1 ml of H<sub>2</sub>O, 1 ml acetone, 1 ml of ethanol and dried under vacuum. PHB concentrations were then determined spectrophotometrically as described by Law and Slepecky, <u>J. Bacteriol.</u>

30

82, 33-36 (1961) using a standard curve and expressed as mg PHB/mg protein.

Plasmid Constructions and Complementation Analysis. Plasmids pLA29, pLA40, pLA41 and pLA42 were constructed by cloning restriction fragments of the pAeT29 insert into 5 the broad host range vector pLAFR3 for complementation analysis of the PHB-negative A. eutrophus strains. pLAFR3 is a derivative of pLAFR1, described by Friedman, et al., Gene 18, 289-296 (1982), containing a pUC8 polylinker cloning site inserted into the EcoR1 site. Different fragments of pAeT9 were cloned into pLAFR3. 10 pLA29 was constructed by ligating the entire 15 kb EcoR1 insert from pAeT29 into the EcoR1 site of pLAFR3. To facilitate the construction of pLA40, pLA41 and pLA42, the corresponding fragments were first cloned into pUC18 to produce plasmids pAeT40, pAeT41 and pAeT42. The fragments were then excised by digestion 15 with BamHI and EcoR1 from the pUC18 plasmids, purified following agarose gel electrophoresis and ligated into BamH1/HindIII digested pLAFR3. To construct pAeT40, pAeT29 DNA was digested to completion with Nde1 and the cohesive ends filled in using the Klenow fragment of DNA polymerase. After separating the fragments 20 on an agarose gel, the 7 kb fragment of interest was purified by electroelution, ligated into the Sma1 site of pUC18 and the recombinant plasmid pAeT40 subsequently identified by restriction analysis after transforming E. coli DH5a cells. This construction eliminates the acetoacetyl-CoA reductase activity since one of the Nde1 sites is located within the structural gene for this enzyme. For the construction of pAeT41, Sma1/EcoR1 digested pAeT29 DNA was separated on an agarose gel and the 5 kb Sma1/EcoR1 fragment purified and ligated into Sma1/EcoR1 digested pUC18 to give the correct plasmid. Deletion of the 2.3 kb Pst1 fragment containing the

B-ketothiolase and acetoacetyl-CoA reductase structural genes by

partial Pst1 digestion of pAeT41 DNA and religation was used to construct pAeT42.

Hybridization Mapping of Tn5 Insertions.

A library of 10<sup>s</sup> individual Tn5 insertion mutants of A. eutrophus 11599 S1 was constructed and 32 potentially PHB-negative 5 colonies, identified by their opaque colony phenotype on nitrogen limited minimal plates, as described above. These were further characterized using Southern DNA hybridization analysis. For the DNA hybridization studies restriction digested chromosomal DNA from each strain was analyzed using both a Tn5 DNA probe (plasmid pRK602) and two plasmids, pAeT10 and pAeT29 which contain the A. eutrophus phbA-phbB locus (Figure 3).

The 32 "opaque" strains represented multiple copies of only three distinct mutant types. These three distinct mutant types are represented by strains PHB #2, PHB #3 and PHB #19. For strains 15 PHB #2 and PHB #3, the transposon Tn5 is inserted into chromosomal Pst1 fragments of 2.3 kb and 0.6 kb, respectively. Both of these chromosomal Pst1 fragments are located on the 15 kb of A. eutrophus DNA cloned in plasmid pAeT29, but not in the phbA-phbB 20 structural genes. Strain PHB #19 has Tn5 inserted into a Pst1 fragment, not present on the pAeT29 plasmid.

More detailed DNA hybridization experiments were performed on the chromosomal DNA from strain PHB #2 and strain PHB #3 to map the site of the Tn5 insertion in each of these 25 mutants. Chromosomal DNA from each of these strains as well as the wild type strain H16 and strain PHB #19 was digested with Sal1, Sma1 and BglII, transferred bidirectionally to nitrocellulose filters and hybridized with Tn5 DNA (pRK602) and pAeT29 DNA probes to map the location of the Tn5 insertions in strains PHB #2 and PHB 30 #3.

The results of a biochemical analysis of wild type H16 and each of the mutuant strains PHB #2, PHB #3, and PHB #19 is

10

15

20

25

30

presented in Table 2. 100 ml stationary phase cultures of each strain were harvested, lysed and assayed for PHB content and Bketothiolase, NADP-specific acetoacetyl-CoA reductase and PHB polymerase activities. Under these growth conditions wild type H16 produces a significant level of PHB (1.3 mg PHB/mg protein, Table 2) and has a high level of all three enzyme activities. Mutant strains PHB #2 and PHB #3 produce essentially no PHB and strain PHB #19 produces only 5% of the wild type level (Table 2). PHB polymerase activity could not be detected in any of these mutant strains, however, the presence of PHB in the lysate of strain PHB #19 indicates that the enzyme is there although the activity is probably below the detection level of the assay. B-ketothiolase activities in all three mutants are reduced to the order of 45% (PHB #2) to 38% (PHB #19) that of wild type strain H16. Similarly, NADP-specific acetoacetyl-CoA reductase activities are around 50% of the wild type level. It was concluded from these data that the PHB-polymerase gene was located upstream from phbA-phbB and that the expression of the latter genes is affected by the Tn5 insertion upstream in the case of strains PHB #2 and PHB #3.

A series of plasmids containing fragments of the A. eutrophus insert of plasmid pAeT29 were constructed in the broad host range vector pLAFR3 for complementation analysis of the PHB-negative mutants.

Each of the recombinant plasmids, pLA29, pLA40, pLA41 and pLA42 were introduced into each of the A. eutrophus strains by conjugation and the resulting transconjugants analyzed on nitrogen limited plates for the restoration of the white (PHB plus) phenotype. Plasmids pLA29, pLA40, pLA41 and pLA42, each of which contains the region upstream from phbA-phbB into which Tn5 has inserted in the chromosome of strains PHB #2 and PHB #3, complemented the mutation in each of these two strains, restoring the white colony phenotype. All four recombinant plasmids also restored the wild type

15

20

colony phenotype to mutant strain PHB #19. In the case of this strain, the Tn5 insertion is located outside the region of the A. eutrophus chromosome contained in each of the plasmids. Control experiments using the vector pLAFR3 resulted in the opaque colony phenotype when introduced into each of the three mutant strains.

Biochemical analysis of each of the complemented strains was performed as described for the characterization of the mutants and these results are also presented in Table 2. The introduction of pLA29 into each of the mutant strains results in the restoration of PHB polymerase activity and PHB biosynthesis (Table 2). In addition, 10 an approximately three to five-fold increase in the levels of Bketothiolase and NADP-specific acetoacetyl-CoA reductase activities was observed. Plasmid pLA40 and pLA41 also restore PHBpolymerase and PHB production to strains PHB #2 (Table 2), PHB #3 and PHB #19, although in the case of pLA40, the phbB gene was disrupted during the construction of this plasmid. Finally, plasmid pLA42 restores PHB polymerase activity and PHB production to all three mutant strains although the phbA-phbB genes have been deleted. In the case of strains containing this plasmid the Bketothiolase and NADP-specific acetoacetyl-CoA reductase activities remain at the same level as the mutant strains (Table 2).

Table 2. Biochemical analysis of mutant and complemented A. eutrophus H16 strains.

Strain	PHB'	Thiolase <sup>2</sup>	Reductase <sup>2</sup>	Polymerase <sup>3</sup>
H16	1.3	8.9	1.0	3.9
PHB#2	< 0.01	4.0	0.5	ND
PHB#3	< 0.01	3.4	0.5	ND
Tn5#19	0.6	3.2	0.4	ND
H16/pLA29	1.0	28.7	9.2	5.3
PHB#2/pLA29	1.5	27.5	3.5	3.8
PHB#3/pLA29	0.9	24.8	4.4	0.7
PHB#19/pLA29	1.8	26.7	4.9	1.0
PHB#2/pLA40	0.9	20.4	0.45	0.6
PHB#2/pLA41		18.0	3.7	0.9
PHB#2/pLA42	1.2	2.0	0.3	4.3
HB#3/pLA42	0.9	5.5	0.5	0.6
HB#19/pLA42	1.2	5.5	0.4	0.6

<sup>1.</sup> mg/mg of protein

2. units/mg of protein

Results shown are the average of two or more experiments.

As described above, the <a href="https://phbB">phbA-phbB</a> genes located on plasmid pAeT29 were expressed in E. coli under the control of the A. eutrophus promoter. Identification of the <a href="https://phbB">phbC</a> gene upstream from <a href="phbA-phbB">phbA-phbB</a> together with the observed decrease in thiolase and

<sup>3.</sup> cpm/min/mg of protein ND: no detectable activity

10

15

20

25

30

reductase enzyme activities in strains PHB #2 and PHB #3 indicates that all three genes are expressed from a single promoter located upstream from phbC. To study this, cultures of E. coli strains containing plasmids pAeT41 and pAeT42 were grown under nitrogen limiting conditions until cells reached stationary phase at which point the cells were harvested, lysed and analyzed. E. coli containing pUC18 was used as a control. The results of B-ketothiolase, acetoacetyl-CoA reductase, PHB polymerase and PHB concentration assays, shown in Table 3, indicate that the lysate of E. coli containing plasmid pAeT41 has a significant level of each enzyme activity and PHB production.

Maxi-cell analysis of the E. coli strains described above was used to determine the molecular weight of the polypeptides encoded by plasmids pAeT41 and pAeT42. Plasmid pAeT10 was included in the analysis as this plasmid expresses the A. eutrophus phbA-phbB genes from the pUC8 vector lacZ promoter. Additional protein bands are present of Mr 40,000 and Mr 26,000 in samples containing plasmid pAeT10 and pAeT41, respectively. Both of these plasmids express the phbA-phbB genes encoding B-ketothiolase (Mr 41,000) and NADPspecific acetoacetyl-CoA reductase (Mr 26,000). Neither of these two proteins is present in the extract of cells containing plasmid pAeT42 which does not contain the phbA-phbB genes. Control experiments in which the vector pUC8 was used gave no signal at Mr 41,000 or Mr 26,000. Both plasmids, pAeT41 and pAeT42, express the PHB polymerase (phbC) gene in E. coli, and in samples which contain extracts of cells containing these plasmids a signal at M, 58,000 is clearly evident. Again, this protein is absent from sample which contains the extract from cells containing plasmid pAeT10 which does not contain the phbC gene and also from a control sample of pUC8 containing cells. An additional band of around Mr 30,000 is present in all samples and was also found in control experiments of cell extracts containing pUC8 and is presumably a vector protein. From

10

15

20

these data it was concluded that the <u>phb</u>C gene expressed in *E. coli* encodes a polypeptide of approximately M, 58,000.

Table 3. Biochemical analysis of recombinant E. coli strains to determine expression of phbC-A-B in E. coli.

Plasmid	Thio- lase U/mg Protein	Reduc- tase U/mg Protein	Polymerase cpm/min/mg Protein	PHB mg/mg Protein
pUC18	0.5	ND	ND	0.015
pAeT41	59.0	2.5	2.4x10 <sup>4</sup>	2.977
pAeT42	0.9	ND	0.02x10 <sup>4</sup>	0.011

ND: no detectable activity

Results shown are the average of two or more experiments.

## Nucleotide Sequence Analysis of phbC.

The 2 kb Smal-Pst1 A. eutrophus chromosomal DNA fragment cloned in plasmid pAeT42 contains the entire structural gene for phbC and probably the regulatory sequences. This fragment was sequenced from both DNA strands multiple times using the dideoxy sequencing method as described above. A single long open reading frame extends from nucleotide 820 to a TGA stop codon at nucleotide 2608. Potential translation initiation codons are present at position 842 (ATG), 1067 (ATG) and 1097 (ATG). Translation from each of these potential start sites would produce proteins of Mr 63,940, Mr 55,513 and Mr 54,483, respectively. Significant amino acid sequence homology between the translation product from the ATG at position 842 to the ATG at position 1067 and the P. oleovarans PHA polymerase gene product, described below, indicates that the first ATG (position 842) is probably correct. Figure 4 presents the entire nucleotide sequence of this region from the Sma1 site to the first 30

20

nucleotides of the phbA he located down tream. The translation product of the open reaching frame from the ATG at position 842 to the TGA at position 2609 is also shown. The PHB polymerase encoded by the phbC gene in plasmid pAeT42 is a polypeptide of 589 amino acids with an Mr = 63,940. The N-terminal 10 amino acids of the phbA gene product are also presented in Figure 4. Additional features of the nucleotide sequence presented in Figure 4 include the C-terminus of an open reading frame which begins upstream from the Sma1 site and terminates at the TGA stop codon at position 76.

Located 85 bp downstream from the <a href="https://phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.com/phbc.c

The expression of phbC alone in E. coli produces neither PHB nor significant levels of PHB polymerase activity (plasmid pAeT42). E. coli appears incapable of synthesizing D-(-)-hydroxybutyryl-CoA, as substrate for PHB polymerase, in the absence of the A. eutrophus phbA-phbB genes. Since the insert of pAeT42 contains both the promoter and structural gene for phbC (plasmid pLA42 complements all PHB-negative mutants, Table 2), it can be concluded that in the absence of available substrate, PHB polymerase is inactive or degraded in E. coli.

The nucleotide sequence of the A. eutrophus chromosomal

DNA insert in plasmid pLA42 encoding PHB polymerase predicts a single polypeptide of Mr 63,940 (Figure 4). Although PHB polymerase has not previously been purified and characterized, the results of E. coli maxi-cell studies indicate a Mr = 58,000 for this polypeptide, in reasonable agreement with that predicted from the gene sequence.

For a number of years, it was proposed that the polymerization of (D)-B-hydroxybutyryl-CoA involves a membrane

bound polymerase which forms a type of barrier between the aqueous environment of the cytoplasm and the hydrophobic crystalline PHB granules. The hydropathy profile of the PHB polymerase polypeptide does not indicate a typical membrane spanning structure. In addition, NMR studies of native PHB granules in Methylobacterius indicate that 5 these granules are in a mobile, as opposed to a highly crystalline solid, state. Together these data lend credence to the idea that PHB biosynthesis does not in fact require a complex membrane bound polymerization system. The mechanism for PHB polymerase proposed 10 in the literature involves two partial reactions. The initial acyl-Senzyme intermediate formation is followed by transfer to a primer acceptor in the second reaction. The predicted primary structure of PHB polymerase has 5 cysteine residues, Cys<sub>246</sub>, Cys<sub>319</sub>, Cys<sub>362</sub>, Cys<sub>436</sub> and Cys45g.

15 Identification of the P. oleovarans PHA Polymerase gene.

The genes involved in the biosynthesis of polyhydroxyalkanoate (PHA) polyesters in *Pseudomonas oleovarans* were also isolated, as follows.

In 1983, de Smet, et al. J. Bacteriol. 154,870-878, identified a polymer produced by Pseudomonas oleovarans TF4-1L (ATCC 20 29347) as poly-B-hydroxyoctanoate. Subsequent studies showed that P. oleovarans could produce a range of PHA biopolymers depending on the carbon source used, i.e., n-alkanes and 1-alkenes (Lageveen, et al., Appl. Environ. Microbiol. 54,2924-2932 (1988)) or fatty acids (Brandl, et al., App. Environ. Microbiol. 54,1977-1982 (1988). The 25 pathway appears to involve the conversion of the alkanes/alkenes to the fatty acid which then enters the fatty acid B-oxidation pathway, resulting in the formation of the D isomer of the B-hydroxyacyl-CoA, which is incorporated into the polymer by PHA polymerase. P. oleovarans has not been shown to incorporate B-hydroxybutyrate 30 indicating that 1) it does not possess the thiolase/reductase enzymes, or 2) the PHA polymerase cannot use B-hydroxybutyrate as a

substrate. The broad range of substrates used by the *P. oleovarans* PHA polymerase make the gene encoding this enzyme particularly interesting for biopolymer engineering of polyesters.

The approach used for isolating the A. eutrophus Bketothiolase and NADP-specific acetoacetyl-CoA reductase using the 5 Z. ramigera B-ketothiolase gene as a DNA hybridization probe was followed, as described above, to isolate the P. oleovarans PHA polymerase gene. Southern DNA hybridization of P. oleovarans chromosomal DNA identified a 6 kb EcoR1 restriction fragment with 10 strong homology to the A. eutrophus PHB polymerase gene (phbC). The 6 kb EcoR1 fragment was cloned in the E. coli plasmid vector, pUC18, by standard procedures to give plasmid pP023. The region which hybridized to the A. eutrophus phbC gene is located as indicated on Figure 5. Nucleotide sequence analysis of the complete 6 kb 15 fragment identified three potential protein coding regions (open reading frames, ORF1, ORF2 and ORF3, indicated on Figure 5). ORF1 begins at the ATG initiation codon nucleotide 554 and ends at the TGA stop codon nucleotide 2231 (Figure 6). This open reading frame is contained in the region of the pP023 insert which hybridizes 20 with the A. eutrophus phbC gene.

ORF1 encodes a polypeptide of 559 amino acids with an M<sub>r</sub> = 62,300 daltons. A comparison of the protein sequence predicted by translation of ORF1 with the amino acid sequence of the A. eutrophus PHB polymerase using the program ALIGN revealed 52% identity between the two proteins. ORF3 begins at the ATG position 2297 and ends at the TAA position 3146, and encodes a protein of 577 amino acids with an M<sub>r</sub> = 64,400 daltons. The predicted polypeptide products of ORF1 and ORF3 have 40% and 42% absolute identity with the phbC gene product. ORF1 complements every PHA negative strain which could be isolated and tested, confirming that this gene encodes the PHA polymerase normally expressed. It is believed ORF3 also encodes a PHA polymerase.

10

15

20

25

However, gene expression studies indicate that ORF3 is not normally expressed.

ORF2 begins at the ATG position 3217 and ends at the TGA position 4948, and encodes a protein of 283 amino acids with an  $M_r = 31,400$  daltons. The gene is believed to encode a PHA depolymerase.

Synthesis of PHB, PHA and similar polymers.

It was established above that it is possible to construct new PHB production strains by introducing the A. eutrophus PHB biosynthetic genes into E. coli resulting in the accumulation of up to 50% dry cell weight as PHB. The construction of new or improved polyester production strains is now possible by the expression of either the PHB biosynthetic genes from A. eutrophus or the PHA polymerase gene and ORF2 and ORF3 from P. oleovarans in a number of systems. Plasmids can be constructed which express the A. eutrophus B-ketothiolase and NADP-specific acetoacetyl-CoA reductase genes in P. oleovarans under the control of the xylS promoter in the broad host range expression plasmids pNM185 (Mermod, et al., J.Bacteriol. 167, 447-454 (1986), or pERD20/pERD21 (Ramos, et al, FEBS LETTERS 226, 241-246 (1986). Alternatively, the broad host range tac promoter expression vectors pMMB24/pMMB22 (Bagdasarian, et al, Gene 26, 273-282 (1983)). These same vectors can also be used to express the A. eutrophus PHB polymerase gene or the three P. oleovarans genes cloned in plasmid pP023 (Figure 5).

Two plasmids, pLAP1 and pLAP2, have been constructed which should express the *P. oleovarans* PHA polymerase gene and ORF2 (pLAP1) or the PHA polymerase gene and ORF2 plus ORF3 (pLAP2) under the control of the *A. eutrophus* phbC promoter (Figure 4).

To construct these plasmids, the 810 bp Sma 1-BstB 1 restriction fragment spanning the first 810 nucleotides of the A. eutrophus insert in plasmid pAeT42 containing the phbC promoter was

ligated into the unique Sma 1 site in the E. coli vector pUC19 to obtain plasmid pAeTB1. The P. oleovarans PHA polymerase gene promoter was removed by using the exonuclease Bal31 to delete 170 bp from the end of Fsp1 digested pPO23 DNA and recovering the PHA polymerase structural gene plus ORF2 by subsequently digesting 5 with Cla1, and cloning the fragment into Sma1/Cla1 digested pBLSK+ vector (Stratagene, La Jolla, CA 92037). The insert of pPOB10 was identified as containing the last 19 bp upstream from the ATG start codon of the PHA polymerase structural gene (nucleotide 535, Figure 6) the complete PHA polymerase structural gene and 10 ORF2. The structural gene and ORF2 could then be recovered on a 2.6 kb BamH1- Xho1 fragment and ligated into BamH1/ Sal1 digested pAeTB1 to obtain plasmid pAeP1. The entire insert of pAeP1 containing the phbC promoter -PHA polymerase structural gene -ORF2 construct was then excised as a 3.4 EcoR1 - Hind111 fragment 15 and cloned into the polylinker region of the broad host range vector pLAFR3 for conjugation into A. eutrophus strains. To construct pAeP2, the 2.6 kb BamH1 - Cla1 fragment from pPOB10 and the 2.5 kb Cla1 - Xho1 fragment from pP023, containing ORF3, were ligated with BamH1/Sal1 digested pAeTB1 to obtain pAeP2. Again the 20 phbC - PHA polymerase - ORF2 - ORF3 construct could be excised as a 5.9 kb EcoR1 - HindIII fragment and cloned into the polylinker region of pLAFR3 to obtain pLAP2. pLAP1 should express both the PHA polymerase and ORF2 in A. eutrophus and pLAP2 should in 25 addition express ORF3. If the genes are not expressed, they can be inserted into the broad host range expression vectors described above for the B-ketothiolase and NADP-specific acetoacetyl-CoA reductase genes.

The PHB polymerase and PHA polymerase genes should prove invaluable for the production of PHA polymers in bacterial and plant systems. However, the cloned and characterized genes can be further modified by constructing fusions of the two polymerases or by

10

15

20

25

30

chemical mutagenesis. Functional polymerase enzymes could then be selected in an appropriate organism by the accumulation of polymer detectable by phenotypic appearance or increased density. This is a straightforward approach to altering the enzyme's specificity to create novel polymerases.

Modification of Polymer Synthesis by Varying Levels of Enzyme Expression.

After isolation and characterization of the polymer genes and gene products from a variety of organisms, as demonstrated for Z. ramigera, A. eutrophus, N. salmonicolor, and P. oleovarans, a means for controlling the expression of the gene products can be established. Overproduction of the Zoogloea thiolase gene was demonstrated by the studies used to define the transcription start site and promoter of the Z. ramigera. Overproduction enables the purification of the enzymes to homogeneity and provides reagent type quantities for analysis and comparison of substrate specificities. In addition, the purified enzymes can be used to synthesize stereospecific substrates for in vitro polymer synthesis. Further, once the transcriptional regulatory mechanism responsible for polymer overproduction is elucidated under a variety of environmental conditions, in vitro systems for the enzymatic synthesis of known polymers, and novel polymers, can be developed to provide new materials. The new materials are then analyzed for chemical composition, molecular weight and rheological characteristics so that maximum utilization can be made of the polymers.

An overproduction system for the Z. ramigera thiolase in E. coli was constructed using the synthetic tac promoter to produce a series of thiolase expression plasmids, the optimum construct in induced in E. coli cells yielding about 20-30% of the total soluble cell protein as thiolase. This method yields thiolase in reagent type quantities, an average of 150 mg of pure thiolase from 1 liter of culture.

20

There are essentially two conditions where gene regulation in Z. ramigera and A. eutrophus may be expected to occur: when carbon starved cells under nutrient limiting conditions are subsequently presented with a carbon source and when cells grown under nutrient limiting conditions have accumulated large amounts of PHB and the nutrient limitation is removed, resulting in PHB degradation.

Transcriptional regulation of the polymer biosynthetic genes is determined as follows. Cultures grown under various 10 conditions are harvested both for enzyme assays (thiolase, reductase and synthetase) and for RNA purification. RNA samples are analyzed in a series of Northern hybridization experiments using the cloned genes as probes. Useful RNA hybridization methodology includes glyoxylation of RNA (McMaster and Carmichael, Proc. Natl. Acad. 15 Sci. USA 74, 4835 (1977)); formaldehyde/agarose gel electrophoresis (Lehrach et al., Proc. Natl. Acad. Sci. USA 16, 4743 (1977)); transfer to nitrocellulose or nylon filters (Thomas, Proc. Natl. Acad. Sci. USA 77, 5201 (1980)); and hybridization with DNA probes labelled with <sup>32</sup>P by nick translation (Rigby et al., J. Mol. Biol. 113, 237 (1977)). DNA probes are prepared from clones pUCDBKI (Z. ramigera); and pAeT3 (A. eutrophus). One result of these studies is the establishment of the operon organization of the genes and the length of the mRNA.

The levels of each of the biosynthetic enzymes are manipulated and the effect this has on polymer synthesis monitored. It is necessary to determine the rate-limiting enzyme(s) in the 25 biosynthetic pathway so that one can increase flux through the pathway by overproducing the rate-limiting enzyme(s); the effect overproduction of each enzyme has on the incorporation of different monomeric units, i.e., the ratio of PHB:PHV in the copolymer produced by A. eutrophus when grown on butyrate; and the result of 30 expression of the genes from one species in other species, for example, the expression of Zoogloea genes in A. eutrophus, and vice

versa, as well as other isolated and characterized heterologous genes from other organisms, e.g., Nocardia and P. oleovarans in Zoogloea and A. eutrophus.

To accomplish overproduction of polymer biosynthetic 5 genes in multiple host organisms, one must use broad host range expression vectors which function in these bacteria. In one instance, enzyme overproduction via gene dosage is carried out. For example, the entire pUCDBKI insert containing the promoter region can be cloned into the vector pSUP104 (Simon et al., Molecular Genetics of the Bacteria-Plant Interaction, A. Pobler, ed. (Spring-Verlag, NY 10 1983) and used to transform Z. ramigera I-16-M. The extent of overproduction of each enzyme is monitored by enzyme assays. A similar approach can be taken for any number of other genes; for example, thiolase; thiolase and reductase; reductase; reductase and 15 synthetase, etc. Secondly, genes can be placed under the transcriptional control of high efficiency promoters, i.e., tac (Gill et al., J. Bact. 167, 611-615 (1986) and tol (Mermod et al., J. Bact. 167, 447-454 (1986). In this case, the constructs are conjugated into mutants defective in the corresponding gene. The expression of the polymer biosynthetic gene or genes of interest can then be tightly regulated, as 20 determined using enzyme assays to monitor the level of overproduction. As each construct is tested, one can begin to monitor the effect on polymer synthesis in a routine manner i.e., the rate and level of synthesis.

25 Modification of Polymer Synthesis by Altering Available Substrate or Enzyme Specificity.

Factors which determine the molecular weight of the PHB or PHA produced by different bacteria can be elucidated by analyzing the molecular weight distribution of the polymers produced by various bacteria. There is little doubt that a number of PHB-producing microorganisms have the ability to incorporate monomers other than D(-)-hydroxybutyrate into the polymer chain. For the PHB-PHV

10

15

20

25

30

copolymer pro iced by A. eutrophus, it has been proposed that propionate is first converted to propionyl-CoA which then acts as a substrate for beta-ketothiolase. The high yields of pure enzymes available from overproduction systems is necessary to determine the range of alternate substrates which each of the three PHB-biosynthetic enzymes can utilize and the types of new PHB-like polymers that can be synthesized in an in vitro system where the available substrates can be strictly controlled.

Although the thiolase and reductase enzymes are an essential part of the biosynthesis of PHB and PHB-like polymers, it is the PHB polymerase which ultimately defines the new types of polymer which can be made. This is facilitated by the development of an *in vitro* system using the enzyme to test a whole range of substrates, many of which cannot enter the cell and therefore cannot be tested for incorporation into PHB by a fermentation process.

Overproduction and purification of more than one reductase enzyme provides a means for comparing the kinetics and specificity of the enzymes. The Zoogloea reductase has been reported to be NADP-specific, however, the A. eutrophus enzyme apparently will use either NAD or NADP. The stereospecificity of this enzyme may make it a useful reagent for the synthesis of D-substrates for PHB polymerase studies. Among the acetoacetyl derivatives to be tested are the oxoester of CoA and oxopantetheine pivaloate (OPP) and the methylene analogs. The ketone but not the oxoester of the methylene analogs is cleaved by Zoogloea thiolase.

Various longer chain alkyl derivatives, in particular the  $C_s$ - $C_s$  linear 3-oxo thiolesters, oxoesters and methylene ketones, may also be useful as substrates for the PHB polymerase, given the existence of  $C_s$ - $C_s$ -betahydroxyalkanoates in B. megaterium, as well as olefins, alcohols and epoxides.

In crude extracts of Z. ramigera, D-betahydroxybutyryl CoA, but not L-hydroxybutyryl CoA, is a substrate for PHB polymerase. It

is expected that other D-hydroxyacyl CoA species can utilize alternate substrates or cosubstrates such as D-betahydroxyvaleryl CoA (HV-CoA). [2-3H]HB-CoA and beta[3-14C]-HV-CoA, each readily preparable by enzymic or chemical synthesis, can be used as substrates and to monitor 3H and 14C content and ratios in polymers precipitated or separated by gel filtration. It is predicted that block copolymer regions, e.g., (HB)<sub>500</sub>(HV)<sub>1000</sub>(HB)<sub>500</sub> can be constructed by careful control of substrate ratios, and leaving groups in elongation phase, e.g., HB-oxo-CoA and HV-S-CoA monomers.

10

15

5

Additional alternate substrates can be tested including branched chain beta-hydroxyacyl CoAs. Testing cannot be done in whole cells since such compounds are not normally transported into the cells. Alternate substrates can be tested for inhibition of normal ["C]-PHB formation first by incorporation of soluble ["C]-HBCoA into insoluble polymer, then as copolymerization cosubstrates and finally for homopolymerization. Alternate substrates can be assayed for  $K_m$ ,  $V_{max}$  relative to HB-CoA and for polymer size determined by calibrated gel filtration studies.

### Method for Production of PHB on a continuous basis.

20

25

30

PHB is produced and stored in bacteria when they are grown under nutrient limiting conditions, usually nitrogen-limiting conditions (for example, 0.1% nitrogen, depending on the species), although culturing the bacteria under conditions providing limited oxygen, phosphate, or other non-carbon nutrient source will also induce PHB synthesis and storage. For example, Azotobacter beijerinckii, a nitrogen fixing bacteria accumulates up to 70% dry cell weight as PHB when grown on glucose/ammonium salts under limiting oxygen. Increasing the available oxygen leads to a decrease in PHB synthesis and a concomitant reduction in the levels of two of the biosynthetic enzymes. The reduction in enzyme levels is indicative of a regulatory mechanism(s) operating at the genetic level. Nitrogen limitation of the growth of Alcaligenes eutrophus results in yields of up

10

15

20

25

30

to 80% dry cell weight PHB. Similarly, Halobacterium and Pseudomonas sp. increase PHB production under nitrogen limitation.

Under non-limiting conditions, the PHB in organisms that normally produce the PHB is rapidly degraded by degradative enzymes. It is possible to mutate these organisms such that the degradative enzymes are inactive or deleted, hence PHB accumulated during limiting conditions of growth cannot be degraded under non-limiting conditions. In order for these bacteria to resume growth, the PHB will be excreted into the medium. Alternatively, it is possible to introduce the requisite enzymes into an organism which does not metabolize PHB (biosynthesis or degradation), enabling that organism to accumulate large quantities of PHB under limiting conditions, and when conditions are changed to non-limiting, the organism should release the PHB into the medium.

By cycling the limiting and non-limiting conditions, it is possible to accumulate the maximum amount of PHB (based on absorbance of the bacteria, which increases as a function of polymer content), then release the accumulated polymer into the medium by changing the conditions to non-limiting conditions which stimulate replication of the bacteria. The organisms can be cultured in conventional fermentation systems for continuous removal of the polymer containing medium without disruption of the bacteria. Expression in plants and production of PHB and PHA Polymers.

As described above with reference to bacterial expression systems, the genes encoding the thiolase, reductase, and/or the polymerase for PHB or PHA can be expressed in plants of a variety of species to produce the desired polymeric product. The advantages of such a system are immediately apparent, decreasing dependence on petroleum-based plastics, and creating an economically useful crop for plants which can grow on a variety of soils.

The first requirement for plant genetic engineering is a system to deliver the foreign DNA to plant tissue. The most popular

10

15

20

25

30

vectors at this time are the tumour-inducing (Ti) plasmids of Agrobacterium tumefaciens, which deliver the DNA by infection. Plant DNA viruses can also be used as vectors, such as vectors based upon the cauliflower mosaic viruses or the Gemini virus vectors. There are also a number of methods of direct gene transfer to plant cells, including chemically stimulated DNA uptake by protoplasts, electroporation, electroinjection of intact plant cells, liposomemediated transformation of protoplasts, and DNA transformation by direct injection into plants. Chemically stimulated uptake involves incubating protoplasts with donor and carrier DNA in the presence of 13% (w/v) polyethylene glycol in 40 mM CaCl<sub>2</sub>. Post-incubation is carried out whereby the PEG concentrations gradually lowered as the CaCl<sub>2</sub> concentration is gradually raised. Electroporation is the process whereby electrical pulses of high field strength are used to reversibly permeabilize cell membranes to facilitate uptake of large molecules, including DNA. Electroinjection and direct injection have the advantage that they do not require formation of protoplasts first. These methods are known to those skilled in the art. See, for example, the review by C. P. Lichtenstein and S. L. Fuller, "Vectors for the genetic engineering of plants", Genetic Engineering, ed. P. W. J. Rigby, vol. 6, 104-171 (Academic Press Ltd. 1987).

The genes can be introduced into the cytoplasm, mitochondria, or chloroplast, either directly or using targeting sequences. Vectors and targeting sequences and promoters for plants are known to those skilled in the art and are commercially available from Pharmacia-LKB Biotechnology, 800 Centennial Ave., Piscataway, NJ 08854-9932, and Stragene, La Jolla, CA.

Any type of plant which produces a useful carbon substrate can be engineered for polymer production. As used with reference to production of polymers in plants, "polymer" includes PHB, PHA, and novel carbon-based polymers synthesized from fatty acids using the disclosed polymerases. If the plant does not form the appropriate

10

fatty acids, the thiolase and reductase genes can be introduced into the plant along with one or more polymerases. The A. eutrophus polymerase polymerizes C4 and C5 substrates. The P. oleovarans polymerase acts on longer substrates, such as C6 to C18 fatty acids, but not short chain fatty acids. The plants can also be modified, preferably by mutagenesis, to block the glycerol ester and fatty acid degradation pathways so that the plant forms the appropriate substrate.

The genes can be introduced into any type of plant. Cereal plants are preferred, such as corn, wheat and rice, since they are widely grown and their genetic systems are well characterized. Other useful agronomic plants include tobacco and high oil seed plants, especially those varieties which grow in desert or in mineralized soil.

The genes can also be introduced into plant cell culture

systems, many of which are known to those skilled in the art. Cell

culture of a variety of cereal and other agricultural crops is described, in detail in <a href="Handbook of Plant Cell Culture">Handbook of Plant Cell Culture</a> vol. 4 edited by D.A.

Evans, W.R. Sharp, and P.V. Ammirato (Macmillan Publishing Co.

NY 1986). A specific example of a plant system in which much

genetic work has been conducted is \*Arabidopsis thaliana\*. Polymer production in cell culture can be manipulated not only by introduction of the cloned genes but also be variation in substrates and culture conditions, as described with reference to production in bacteria.

Modifications and variations of the present invention, a

method for making polyhydroxybutyrate and polyhydroxybutyrate-like
polymers having carbon-carbon backbones using recombinant
engineering according to the foregoing detailed description, and the
resulting polymers, will be obvious to those skilled in the art. Such
variations and modifications are intended to come within the scope of
the appended claims.

We claim.

1. A method for constructing polyester biopolymers in a host comprising:

selecting a host for expression of genes encoding enzymes required for synthesis of polyhydroxyalkanoates,

introducing into the host isolated structural genes encoding enzymes selected from the group consisting of beta-ketothiolases, acetoacetyl-CoA reductases, polyhydroxybutyrate polymerases, and polyhydroxyalkanoate polymerases in combination with regulatory sequences for expression of the genes in the host,

expressing the enzymes encoded by the introduced genes, and

providing appropriate substrates for the expressed enzymes to synthesize polyhydroxyalkanoates.

- 2. The method of claim 1 wherein the polyhydroxyalkanoate is polyhydroxybutyrate.
- 3. The method of claim 1 further comprising selecting the enzymes on the basis of their substrate specificity.
- 4. The method of claim 1 further comprising altering the substrate specificity of the enzymes by modifying the genes encoding the enzymes.
- 5. The method of claim 1 further comprising providing regulatory sequences in the expression vector which control expression of the genes encoding the enzymes required for polymer synthesis in response to specific inducers.
- 6. The method of claim 5 wherein the inducer is selected from the group consisting of temperature changes and specific substrate.
- 7. The method of claim 1 wherein the genes are from bacteria selected from the group consisting of Zoogloea, Azotobacter, Alcaligenes, Bacillus, Nocardia, Clostridium, Halobacterium, Escherichia, Pseudomonas and Rhodospirillium.

- 8. The method of claim 1 further comprising expressing the genes in bacterial hosts which are deficient in at least one enzyme required for polymer synthesis.
- 9. The method of claim 1 further comprising expressing the genes in plant hosts.
- 10. The method of claim 1 wherein the genes encode enzymes stereospecific for the D isomer of the Hydroxyacyl-CoA substrate.
  - 11. The biopolymer produced by the method of claim 1.
- 12. An isolated DNA sequence hybridizing to a gene encoding polyhydroxybutyrate polymerase at a temperature of 65°C in a NaCl, sodium citrate, sodium phosphate solution for a period of approximately 16 to 18 hours.
- 13. The sequence of claim 12 encoding an enzyme expressed in a bacteria selected from the group consisting of Zoogloea, Azotobacter, Alcaligenes, Bacillus, Nocardia, Clostridium, Halobacterium, Escherichia, Pseudomonas and Rhodospirillium.
  - 14. The sequence of claim 13 as shown in Figure 4.
- 15. An isolated DNA sequence hybridizing under standard conditions to a gene encoding polyhydroxyalkanoate polymerase.
- 16. The sequence of claim 15 encoding an enzyme found in Pseudomonas oleovarans.
- 17. The sequence of claim 16 having the sequence shown in Figure 6.
- 18. The sequence of claim 16 having the sequence shown in Figure 6 as open reading frame 1, nucleotides 554 to 2231, and open reading frame 3, nucleotides 3217 to 4948.
- 19. An isolated DNA sequence having sequence shown in Figure 6, open reading frame 2, nucleotides 2297 to 3146, or a sequence hybridizing thereto under standard conditions.
- 20. Polyhydroxybutyrate polymerase expressed from an isolated structural gene.

- 21. The polyhydroxybutyrate polymerase of claim 20 encoded by a gene found in a bacteria selected from the group consisting of Zoogloea ramigera and Alcaligenes eutrophus.
- 22. Polyhydroxyalkanoate polymerase expressed from an isolated structural gene.
- 23. The polyhydroxyalkanoate polymerase of claim 22 encoded by a gene found in *Pseudomonas oleovarans*.
- 24. The protein expressed from open reading frame 2, nucleotides 2297 to 3146, of Figure 6.
- 25. A system for synthesizing biopolymers having polyester backbones comprising
- a host for expression of genes encoding enzymes required for synthesis of polyhydroxyalkanoates,

isolated structural genes encoding enzymes selected from the group consisting of beta-ketothiolases, acetoacetyl-CoA reductases, polyhydroxybutyrate polymerases, and polyhydroxyalkanoate polymerases in combination with regulatory sequences for expression of the genes in the host, and

appropriate substrates for the enzymes to synthesize polyhydroxyalkanoates.

- 26. The system of claim 25 wherein the host is selected from the group consisting of bacteria and plants.
- 27. The system of claim 25 wherein the substrates are selected from the group consisting of acetoacetyl derivatives, olefins, esters, alcohols, diols, epoxides, D-hydroxybutyryl CoA, D-beta-hydroxyvaleryl CoA, branched chain beta-hydroxyacyl CoAs, monomeric acyl CoA analogs of D-3-hydroxybutyryl-CoA, acyl thiolester CoAs, and combinations thereof.

# FIGURE 1

Eco RI <u>CAMITG</u>ETCA rece*ce*rer cerettire ittirace ircacera ecocadada Acoereero -130 -170  $\frac{1}{10}$ ιτικοστη  $\frac{1}{10}$ υση ήτητοσλο Αντισσίας της  $\frac{1}{10}$  το ποποστος αντισσίας το εντισσόμο  $\frac{1}{10}$  το  $\frac{1}{10}$  το  $\frac{1}{10}$ iotoccatoc occadocat detococaae Atocatitic Atocitocos occeeçoca icaaatitia ococeoooo Atateecegi jeasieteei Att the state of t ATG ANG OCC OCC GTG CGG CNG CNG ACC OCC TOC OCC ATG ANG CNG CTT TOC OCC NG Lys Ala Gly Val Pro Gln Glu Ala Thr Ala Trp COV Bot And Gln Lou Cya Cly 310 84 843 525 က် ကို 326 GTC CAA CCC AAC CAA CCC TIC CCC CAG CCC TCC CCC CTC AAC AAC CAC CTC CCC Val Glu Ala Aan Olu Ala Phe Ala Ala Gln Ala Cys Ala Val Aan Lys Asp Lou Gly 1030 CAC ACC Ale Val 35 Glu Ala Ser Arg Arg Gly Ile Gin ATC CCC CAC CCC ACC ACC CAT ICC ATC CCC AAC CTC CCC CCC IIC CAC AAC CAA CCC IIe Arg His Ciy Ala Thr Leu Asp Ser Met Ala Lys Leu Arg Pro Ala Phe Asp Lys Clu Ciy Thr 750 Ne CH AAC ACG Asn Thr ដ្ឋ ğ CIC TOC AIC OOC OOC AIC OOC GIO OCO AIC Leu Cys Ile Gly Gly Gly Het Gly Vel Ale Het Cys Ile ore one one one Are city val Ala city city Mat city ATC ATC AS CAC CCC CTC Het II. Lys Asp Gly Lou 530 OCC CAG AAT GTC CCC AAG CAG TOO CAG CTT ICC COC CAG CAG CAG CAC CAC TTC Ala Glu Asn Val Ala Lya Glu Trp Cln Leu Ser Arg Asp Glu Gln Asp Ala Pheson Arg Lys ACC CCC CCC ATC CCC CCC TCC And Ser ... ככ זכא מבן זאן ככא אא כדן זככ כאן כאכ כדם ככם ככם ככם אכם אפע אפע הפלון בַכאַ בַּבּ Sall ¥: OCC OCC GID AND AND GC CAT INC AAG AND AND CAC ACG AIA GIY VAI LYS HOE GIY ASP Phe LYS HOE ING ASP Thr S10 ATC ACC CAA CCC Met Ser Clu Ala 870 .55° CTC GOS ATG CAG ATG CCC AGG GGC GAT GGC AGG Lau Gly Het Gln Gln Ile Ala Thr Gly Asp Ale Ser 390 GTC. CAT CCC AAG GTC ATG CCC Val Asp Pro Lym Val Hat Gly 950 CLG ATG ANG CCT CCC CCC CCC ANG CCT CTC CCC ACC Clu Hot Lys Arg Arg Cly Ala Arg Lys Cly Leu Ala Thr 84 85 22 .84 48 AAT CAC Asn Asp 524g con the Ave Are con car are cad ony inp Lys ile city Asp Leu Asp 1010 Ala Thr Val IIe Ser Ala Val Leu CCC CCC CCC CAG Pro Ala Arg Gln 290 18 AC Ala Thr 63 53 85 ATC CC CCC CAT TCC CCC CAT Met Ala Pro His Cym Ala His 470 TTC DAG GOX DAG CAC ATG GOX Phe TyT CITY TYT HIS Het CITY \$50 8gg OCC COC ATC CIT ICC TOC CLY Arg Ile Val Ser Irp CAC CAA TAT Asp Clu tyr CCC TCC ATC CTC AAC Pro Ser IIo Val Asn 1090 53 Ala Ser ¥¥ 84 36. 36. 36. 36. 36. 36. Ser CIVI Lau Arg CAT CCC ASP AIA OCC ARC CAG AAC AAG Cln Asn Lys ACASTCCAST 33 52 84 84 R S TT Page 23 85 85 53 38 žg 88 ĘĘ

# FIGUREZ

• : <u>Saii</u> Gicaactcaaaaatcagictagg<u>gaag</u>tgaggacaggagggagggagggaggaacagggggggarcggggggatcgcggggcatcggcgcatt H <u>s r y a l</u> t q a s r q i a a a i TCOATTOCOCTOAAGCCGGCGGATACAAGTGGCTGCCAGCAATGACGATGCGGCCAAGCCCTTCAAGGCGGAAACGGGC s i a l k a a g y k y a a s y a g w d b a a k p f k a e t g CTCGTCAACAATGCGGGCATCACCAAGGACGCGAATGACGCCCGACCAGTGGAATGCGGTCATCAACACCTCACG L y n n a g i t k d a b f h k b t p d q b n a y i h t n l t GOTCTCTTCAACATGACCCATCCGGTCTGGCATGCGCAGCTTCGGCCGCATCGTCAACATCTCCTCGATCAACGGCCAG O L f n m t h p v w s o m r d r s f o r i v m i s s i n o q AAAQQQCCAQATQAQTCAAQCQAACTATTCCQCCGCAAQQCGGACCACCAAAQCGCTQACGAAAAAGCGCGGCAAA K'G G W G R A W Y S A A K A G B L G F T K A L A'G E G A A K ODCATCACOGTCAACOCCATCTOCCCCOOCTATATCOOAAATOCTOCOCCATTCCOOAAAAOOTOCTOAACOAGGOATCATC  ${f a}$  of  ${f i}$  of  ${f i$ coochantecedatedaceaecteageaageaaaceaaktaategeegetteateatetteeteaecteaecteateaecf R a f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f R f ROOCTCOACCATCTCOOCOAACOOCCOOCCATTCTTCOTCTOATACCOOCCACACO<u>OAA</u>CO<u>coccoc</u>ttco<u>ococc</u>ttco<u>occottic</u>at o s t i s a m o o q f f f v . QTGGTATGCTGGTCGAAAQQAGAGCCCCGAATQAAACAGGAAAAGCTGGATGAGGGGGGGATTGCCGAAGGCGCTGGCATCCCTCQAAGQAT **ΦΟΑΤΤΕΩΕΤΕΘΩΕΤΩΑΚΩΩΕΑΤΕΩΕΩΑΤΕΩΑΘΑΑΛΩΕΩΕΤΑΚΑΣΩΤΤΕΑΛΩΛΩΣΕΤ<u>Τ</u>ΕΓΕΩΣΩΑΟΩΕΩΤΤΕΩΩΩΕΤΕΑΤΩΛΕΣΩΑΟΩΩΩΩΩ** ΤΦΟCΘΒCΘΒΑΒΑΑΑΤΤΕΑΑCCACCATCCCΘΑΑΤΘΩΤΤΕΑΑCGTCTACAATCBC<mark>ΘΤζόλτ</mark>ΘΤΘCGGCTDACCAACCAACGATOCGGGGG CAAGOACOAGCACOACOTCAGOAAAGCCGCGTGAGGCGGCGGTTCTGGCCGGCTCTCAAGCGCGATGCGGCAGGTTCCCTTCGCCG TOACCOAGCTOOACTTCAAGCTGGCGGCGGCGATGGAGAAGGCGGCGGCCTTTCGCACOAAGAGTTGAAGCGGCGTGAACGCATTCCCAT ATBATCAOCACCAGACTOTTCCAOCACCACCCCAAQOQAQQCCTTQAATQQACQATGTAAAQATCAOCOAQATTCTOCTACCCGQCOA COATCTCGTCGCGTCCTACTATTGCGCCCCAATCCCCACACGCCGGGGCGCGCGGGGATCCTGCTGGCGCGCTCGCCTATTTTGT CGCAATCGQQGTCAAATTCTTQCCCTATTCCTTGTQCCATTAQQCATCTCQACQQQQTTCATCTCACCAAATQQQAACTQQGCACQCT 491AACCTDAATTAA9TCAAAATAAATCATCATCATCATCATGACATCGACACCOGCGTQAAAQQCATCCGCCTTCGQCCGGGCATCAC TBBCAABACAACCBBCAABAAAACATBTTCBTAABAAABCTCBCAACCBCACTCBCCCTCBTCCTTBCBACCBCAABBCBTTBCCTCBGCB CAAACQCCQACQCQCATTCAQCAQTTCAATQCCTQQQQCQCCTATTCCTATCAQQCQAATQQCQQCAAQQTCTQCTACQTQTGCQTQ Smaj CCOAAGGAAAAQACGCCGCCGGG 

Pst '' <u>croch c</u> errece contrecatre a A a contrea contrea contrates and the contrea contre contre contrates and the contrate cont	Ju Trccattgaaaggac	TACACAÁTGACTG Metthra	ou Accircrcarcer spyalyallleya	/0 ATCCCCCCCCCCC	CCGCGGTCGGCAAG	90 ກາກເດເດດີດ ການເດເດດີດ
011		c	150	6	manada da sa	10017017
TCGCTGGCCAAGATCCCGGCACCGGAACTGGGTGCCGTGGTCATCAAGGCCGCGCTGGAGCGCGCCGCGCTAAGCCGGAAGTGAGCGAAGTGAGCGAAGTC SerlewalalysileProalaProGluLeuGlyAlaYalYalIleLysAlaAlaLeuGluArgAlaGlyYalLysProGluGlnYalSerGluYal	GCACCGGAACTGGGT AlaprogluLeugly	GCCGTGGTCATCA	AGGCCGCGCTGGA YSAlaAlaLeuGl	ro scecececerca sargalaciyyally	Accccaaccaigno ysprogluginyal	190 GCGAAGTC SerGluYal
210	,	230	250		270	
ATCATGGGCCAGGTGCTGACGGCGGTTCGGGCCAGAACCCGGCACGCCGCGGATCAAGGCCGGCC	ACCGCCGGTTCGGGC ThrAlaGlySerGly	CAGAACCCCGCAC GlnAsnProAlaA	GCCAGGCCGCGATU rgGlnAlaAlaIl	PAGGCCGGCCTGC	CGGCGATGGTGCCGC	SCCATGACC AlametThr
290	310	330	••	350	370	
A†CAACAAGGTĠTGCGGCTCGĠGCCTGAAGGĊCGTGATGCTĠGCGGCCAACĠCGATCATGGĊGGGCCACCCĊGAGATCGTGĠTGGCCGGCGCGCAA IleasnlysyalcysGlyserGlyLeuLysAlayalMetLeuAlaAlaAsnAlaIleMetAlaGlyAspAlaGluIleyalyalAlaGlyGlyGly	rcccccccaagcc sergly Leulysala	GTGATGCTĠGCCG ValMetLeuAlaA	ccaaccccarcard laasnalaileHe	GCGGCCGACGCCG	AGATCGTGĠTGGCCC lulleValValAlac	GCGGCCAG
390	410	430		450	470	
GAAAAĆATGAGCGCCCGCCGCACGTGCTGCCGGGČTCGCGGATGGTTTCCGCATGGGCGATGCČAAGCTGGTCĜACACCATGATCGTCGACGGČ GluasnMetSeralaalaProHlsYalLeuProGlySerargAspGlyPheargMetGlyAspAlaLysLeuValaspThrMetIleValAspGly	cccacciccrcccc	GCČTCGCGCGATĠ GlyserArgAspG	strrccccaiges Lypheargmets)	GATGCCAAGCTGG AspAlaLysLeuY	TCĠACACCATGA†CC alaspihimetilev	TCGACGC
490	510		530	550		570
CTGTGGGACĠTGTACAACCÁGTACCACATĠGGCATCACCĠCGGAGAACGŤGGCCAAGGAÁTACGGCATCÁCACGCGAGGĠGCAGGATGAĞTTCGCC LeutrpaspyaltyrasnGlntyrHismetGlyileThralaGluasnyalalaLysGlutyrGlyileThrargGlualaGlnaspGlupheala	agtaccacargggc lntyrhismetgly)	ATCACCĠCCGAGA/ IleThralaGluA:	\CGTGGCCAAGGAi nvalalaLysGlu	TACGGCATCACAC Tyrglyllethra	GCGAGGĆGCAGGATG FGGluAlaGlnAspG	AGTTCCCC luPheAla
290	610	0	630	650	•	670
GTCGCCTCGCAGAACCCGGAAGCCGGCAGAAGCCCGGCAAGTTTGACGAAGAGAATCGTCCGGGTGCTGATCCCGCAGGGCAAGGGCGAACGGCGACCCG ValGlySerGlnAsnLysAlaGluAlaAlaGlnLysAlaGlyLysPheAspGluGluIleValProValLeuIleProGlnArgLysGlyAspPro	cccaaccccccac	naggcgggcangt Lysalaglylyspi	TGACGAAGAGATO eAspGluGluIle	GTCCCGGTGCTGA ValProvalLeuI	icccgcagcgcaagg leProglnArgLysG	GCGACCCG 1yAspPro
069		710	730		750	
GTGGCCTŤCAAGACCGAČGAGTTCGTGČGCCAGGGCGČCACGCTGGAČAGCATGTCCĠGCCTCAAGCČCGCCTTCGAČAAGGCCGGCÅCGGTGACC ValAlaPheLysThrAspGluPheYalArgGlnGlyAlaThrLeuAspSerMetSerGlyLeuLysProAlaPheAspLysAlaGlyThrYalThr	AGTTCGTGĊGCCAGC lupheYalArgGlnC	3GCGCCACGCTGGA 31yAlaThrLeuAs	.cagcargrccigg pserMetsergly	CTCAAGCCCGCCT LeulysProAlaP	rcgačaaggcgggcå neasplysalaglyt	CGGTGACC hrvalthr
770	790	810	83	830	850	
GCGCCAACGCCTCGGGCCTGAACGACGGCGCCGCGCGCGGTGGTGGTGATGTCGGCCGAAGGCCAAGGACTGGGCCTGACCGGCTGGCCAGGCAACACGCAACACGAACACACGAACACGCTGGCCTGGCCTGGCCTGGCCAACGCAACAGGAACAGGAACTGGGCCTGACCGCTGGCCAACACACAAAAAAAA	TGAACGACGCCCCC euAsnAspGlyAlaA	sccccccrccrcr laAlaValValVa	GATGTCGGCGGCC lMetSerAlaAla	AAGGCCAAGGAAC; Lysalalysgluld	rgggccrgacccgc euglyteuThrProt	rgccacg ewlathr
870	890	910		930	950	
Ava. ATCAAGAGCTATGCCÁACGCGGGGGTGGGGTGATGGGGÁTGGGCCGGGGGCCGGGCTCCAAGCGCGCCCTGTGGGGGGGG	ccggrgicgarccca laglyvalaspprol	AĠĠŦĠAŦĠĠĠĊĂŦ ŊsValMetĠlyMe	GGCCCGGiGCCG tGlyProValPro	3CCTCCAAGCGCGC AleserLysArgAl	ictrorcececece abeuserArgAlag	Ava. A <u>cr<u>ccac</u> lutrpthr</u>
970 AVA	066	. ·	1010	1030	1050	20
CCGCAAGACCT <u>GGAC</u> TGATGGAGATCAACGAGGCCTTTGCCGCGGAGGCGCTGGGGGGGG	rccacarcaaccacc ecclulleasnolua	ccrrrccccccc. lapheAlaAlaGl	GGCGCTGGCGGTĠ nAlaLeuAlaYalf	ACCAGCAGÀTGGG IlsGlnGlnMetGl	CTGGGACACCTCCA YTrpAspThrSerLy	AGGTCAAT 7.sYalAsn
1070	1090		1110	1130		1150
GTGÀACGGCGGCGCCATCGCCATCGGCACCGCGCGTCGGCGTGCCGTATCCTGGTGCCTGCTGCTGCACGAGATGAAGCGCCGTGACGCG ValAsnGlyGlyAlaIleAlaIleGlyHlsProIleGlyAlaSerGlyCysArgIleLeuValThrLeuLeuHisGluMetLysArgArgAspAla	CATCGGCCACCCGA LalleGlyHisProi	rcccccccrcccc leclyAlaSerGly	CTGCCGTATCCTGC YCysArgIleLeu\	TGACGCTGCTGCA alThrLeuLeuH1	CGAGATGAAĠCGCC SGlumetlysarga:	TGACCCG

FIGURES (prife 1)

•		
AAGAAGGGCCTGGCCTCGCTGTGCATCGGCGGCGCGTGGGCGTGGCAGTCGAGTCGAGCGCAAATAAGGAAGG	GTCÓAGCGCAAATÁAGGAAGGGGÍTTTCCGGGGCCGCGCGCG Valgluarglys *	• .
1260 Ava Gitggcgc <u>ggac</u> cgcgacgacganaccaatcaaggacatgacatgacatigcgtatgtgacgatgacgaggaigga Methrglmargllæalatyrvalthtglyglyhetglyglyllegly	1340 JAGCGCATÍGCGTATGTGÁCCGGCGGGAÍGGGTGGTATĆGGA JLAK9IleAlaTyrValThrGlyGlyGlyGlyIleGly	
1360 1380 14	1400 1420.	
ACCGCCÁTTIGCCAGCGGCTGGCCAAGGATGGCTTTCGTGGTGGCCGGTTGCGGCCCCAACTCGCCGCGGGGAAAAGTGGCTGGAGCAGCAG Thraibíle CysginargleualalybaspgiypheargyalyalalagiycysgiyproasnSerproargarggiulystrpleugiugingin	GCCCCAACTCGCCGCGCGGAAAAGTGGCTGGAGCAGCAG 11yProAsnSerProArgArgGluLysTrpLeuGluGlnGln	,
1440 1460 1480	1500 1520	
ÁAGGCCCTGGĠCTTCGATTTĊATTGCCTCGĠAAGGCAATGŤGGCTGACTGĠGACTCGACCÀAGACCGCATŤCGACAAGGTCAAGTCCGAGĠTCGGC LysalaleuglypheAsppheIlealaSerGluGlyAsnYalAlaAsptrpAspSerThrLysThralaPheAspLysValLysSerGluValGly	CGACCÀAGACGGCATTCGACAAGGTCAAGTCCGAGGTCGGC erthriysthralapheaspiysvallyssergluvalgly	
1540 . 1560 1580	1600 1620	
GAGGITGATGTGCTĞATCAACAACGCGGTATCACCCGCGACGTGGTGTTCCGCÂAGATGACCCGCGCGGCTGGGATGCGGTGATCGACACCAAC GluyalAspyalLeulleAsnAsnAlaGlyIleThrAfgAspyalyalPheArgLysHetThrArgAlaAspAraAliaYalIleAspThrAsn	AGATGACCCGCCCGACTGGGATGCGGTGÅTCGACACCAAC ysMetThrArgAlaAspTrpAspAlaYalileAspThrAsn	
1640 1660 1680	1700	
CTGACCTCGCTGTTCAACGTCACCAAGCAGGTGATCGACGGCATGGCCGACGTGGCTGGGCGGCCGCATCGTCAAATCTCGTCGTGAACGGGCAG LeuthrSerLeupheAsnValthrLysGlnValileAspGlyMetAlaAspArgGlyTrpGlyArglleValAsnIleSerSerValAsnGlyGln	GCTGGGGCCGCATCGTCAACATCTCGTCGGGAACGGGCAG 1yTrpG1yArg11eYalAsnI1eSerSerYalAsnG1yG1n	
1740 1760 1780	1800 1820	
AAGGGCCAGTTCGGCCAGACCAACTACTCCACCGCCAAGGCCGGCC	TCACCATGCCACTGGCGCAGGAGTGGCGACCAAGGGCGTG heThrMetalaLeualagingluyalalathrLysGlyyal	••
1840 1860 1880	30 1900	•
ACCGTCÁACACGGTCTCTCCGGGCTAŤATCGCCACCÓACATGGTCAÁGGCGATCCGCCAGGACGTGCTCGACAAGAŤGGTGGGGACGATCCCGGTC ThrYalasnThrYalSerProGlyTyrIleAlaThrAspHetYalLysAlaIleArgGlnAspYalLeuAspLysIleValAaThrIleProVal	ocaggaconocasonos de la conocarca de la colocación de la	
1920 1940 1960	1980 2000	
Aggogocorgogocorgogaagaarogocorogaanongaa ka k	.GGAGİCCGGTTICTĞGACCGGCGCGACTTCTCGÖTCAAC .uGluSerGlyPheSerThrGlyAleAspPheSerLeuAsn	
2020 2040 2060	2080 2100	
GGCGĠCCTGCATATĠGGCTGACCTĠCCGGCCTGĠŤTCAACCAGTĊGGCAGCCGGĠGGTGGCGCGCGTATTGCĠGTGCAGCCAĠCGGGGGGCAĊA GlyGlyLeuHlsMetGly *	CTGGCGCCCCCATTGCGGTGCAGCCAGCGGGGGGCACA	
2120 . 2140 2160	2180 2200	
AGGCGGCGGCGTTTCGCTTTCGCCGCTTTCGCGGGCCGTCAAGGCCCGCAATCGTTTCTGCCCGCGGGGGGATTCCTCGCCTTTTTGCGCCAAT	cairrerececececacarreciecerrrricececaar	
2220 2240 2250	2280 2300	
TCACCGGGTTTTCCTTAAGCCCCGTCGCTTTTCTTAGTGCCTTGTTGGGCATAGAATCAGGGCAGCGGCGCAGCCAGC	רכאפספים במפכם באפכב בכב בים בים בים בים בים בים בים בים בים בי	
2320 PSt Georgeographs	•	
מריניו זיין פרינים פרינ		

FIGURES (paye 2)

#### FIGURE 4

rnlc Sma PHB polymerase TAGAGCACCACGCCACCGGCGCCATGCCATACATCAGGAAGGTGGCAACGCCTGCCACCACGTCGTGGTGATCGCCATCATCAGCG 270 361 360 450 TCTCGGAATCGCTGACGATTCCCAGGTTTCTCCGGCAAGCATAGCGCATGGCGTCTCCATGCGAGAATGTCGCGCTTGCCGGATAAAAGG . 541 630 721 CCTCCCGCTTTGGGGGAGGCGCAAGCCGGGTCCATTCGGATAGCATCTCCCCATGCAAAGTGCCGGCCAGGGCAATGCCCGGAGCCGGTT 720 811 CGAATAGTGACGGCAGAGAGACAATCAATCATGGCGACCGGCAAAGGCGCGGCAGCTTCCACGCAGGCAAGGCAAGGCAACCATTCAA
H A T G K G A A A S T Q E G K S Q P F K 810 900 GCTCACGCCGGGGCCATTCGATCCAGCCACATGGCTGGAATGGTCCCGCCACGTGGCAGGCCACTGAAGGCAACGGCCACGCGGCCGCGTC Y T P G P F D P A T W L E W S R Q W Q G T E G N G H A A A S G I P G L D A L A G V K I A P A Q L G D I Q Q R Y H K D F S 1000 **VECECLELECTEGECYLEGCCEVEGECTVEECCEVECCETEGELECTECTECTEGECGCCTLEGCCCTCCV** A L W Q A H A E G X A E A T G P L H D R R F A G D A W R T N 1170 LPYRFAAAFYLLHARALTELADAVEADAKT 1260 1350 1351 GCTGATCGAGTCGGGCGAATCGCTGCCGGCGTGCGCCAACATGATGGAAGACCTGACACGCGGCAAGATCTCGCAGACCGACGA · 1440 LIESGGESLRAGVRNHHEDLTRGKISQTD 1441 GAGCGCGTTTGAGGTCGGCCGCAATGTCGCGGTGACCGAAGGCGCCGTGGTCTTCGAGAACGAGTACTTCCAGCTGTTGCAGTACAAGCC SAFEVGRNVAVTEGAVVFENEYFQLLQYKP 1530 1531 GCTCACCGACAAGGTGCACGCGCGCCGCTGCTGATGGTGCCGCCGGGGACAGTACTACATCCTGGACCTGCAGCCGGAGAGCTC
L T D K V H A R P L L H V P P C I N K Y Y I L D L Q P E S S 1620 GCTGGTGCGCCATGTGGTGGAGCAGGACATACGGTGTTTCTGGTGTCGTGGCGCAATCCGGACGCAGCATGGCACACCAGCACCTGGGA V E Q G H T V F L V S W R N P D A S K A G S T W D 1710 1800 V S T A L A V L A A R G E H P A A S V T L L T T L L 1890 1891 GEYCITLÉGCEVCYCE GECYLCTLCEYCELTIFLEGYCEYCEVC GECCYLELGC GECET GECCET GECC D F A D T G I L D V F V D E G H V Q L R E A T L G G G A G A 1980 CCCGTGCGCCGCCGCGCCTTGAGCTGGCCAATACCTTCTCGTTCTTGCGCCCGAACGACCTGGTGTGGAACTACGTGGTCGACAA
PCALLRGLBLANTPSFLRPNDLVWNYVVDN 2070 2071 CTACCTGAAGGGCAACACGCCGGTGCCGTTCGACCTGCTGTTCTGGAACGGCGACGCCACCAACCTGCCGGGGCCGTGGTACTGCTGGTA
Y L K G N T P Y P F D L L F W N G D A T N L P G P W Y C W Y 2160 2161 CCTGCGCCACACCTACCTGCAGAACGAGCTCAAGGTACCGGGCAAGCTGACCGTGTGCGGCGGTGGACCTGGCCAGCATCGACGT LRHTYLQNELKVPGKLTVCGVPVDLASIDV 2250 PTYIYGSREDHIVPHTAAYASTALLANKLR 2341 CTTCGTGCTGCGTCGCCCATATCGCCGGTGTGATCAACCCCGCCCAAGAACAAGCGCGCACTGCGACTAACGATGCGGTGCC V L G A S G H I A G V I N P P A K N K R S H W T N D A L P 2430 ESPQOWLAGAIEHHGSWWPDWTAWL 2520 CGCGAAACGCGCGCGCGCCCAACTATGGCAATGCGCGCTATCGCGCAATCGAACCCGCGCGTGGGCGATACGTCAAAGCCAAGGCATG 2521 A K R A A P A N Y G N A R Y R A I E P A P G R Y V K A K A • 1611 ACGCTTGCATGAGTGCCGGGGTGCATGCACGGCGCGCGGGGGGGCGTGCCAGGTTCCCTTCCCGTTTCCATTGAAAGGACTACACAATGAC 2700 VIVSAARTAVGKFGGSLAKI

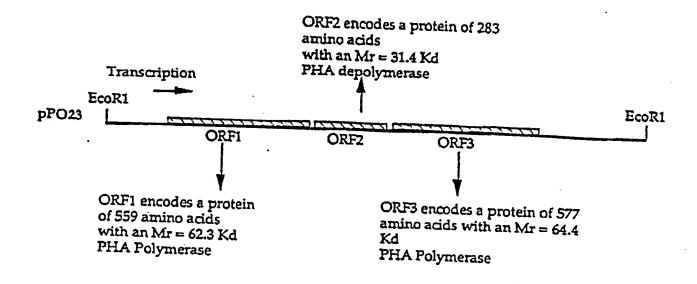


FIGURE 5

# FIGURE 6 (page 1)

1	Eco K1	
_	GAATTCCTGCGCTTGCACTCCCCCTCCGCCGAGGTCCAGGGCCACGGTAACCCTACACTTCCGCAACATCAACGTCGGCCTCAGCGGCCTGGAAC	:
1		
1	CTGCCGGCCAATACGCACTGAAACTGACCTTCGACGACGGCCCATGACAGCGGCCTGTTCACCTGGGAATACCTCGAGCAGCTGTGCCTGCGCCAGGAACA	
ı	GCTGTGGGCCGAGTACTCGAGGACTGCACAAGGCCGGGAAATCCCGGGACCTGCGGGGTGGTCAAACTCATGCTCTAGCGCAAGGCCTGCAGG	
	ATTIAGAGCGCATTITCTAAAATCATCTGTTTGAATGACTTACAGACAGCCCAGTGACGGCCTGTCTTGCGCATTACATGAAAGTCGGGTAACCAATGGG	
L	GGTGGCMAGTTCCCTGCATCAMATTGCAGGTAGTCAGAACCCTTGGCAGCACCGCTGTTCCTTATCACTGGTCACCCGAGTAGCAGTACCGGGCTCAGAAC	
l	TOTOCACCOCCACACCACCACCACCACCACCACCACCACCACC	
	MetSerAsnLysAsnAsnAspGluLeuGlnArgGlnAlaSerGluAs	
	CACCLTGGGGCTGAACCCGTCATCGGTATCCGCCCAAAGACCTGTTGAGCTCGGCACCGTGCTGCGCCAAGCCGTGCGCCAACCGCTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGCTTGCACAGAGACAGAAGACAGAGACAGAGACAGAAGAAGAAG	
	GCCAACCATETTCCCTCCCCCCCCCCCCCCCCCCCCCCCC	
	The state of the s	
	CATGGAGCAACACCCACTITACCGCCGCTACCTGCAAACCTATCTGGCCTGGGGCAAGGAGCTGCAGGACTGGATCGGCAACAGCGACCTGTCGCCCCA laTrpSerAsnAsnProLeuTyrArgArgTyrLeuGlnThrTyrLeuAlaTrpArgLysGluLeuGlnAspTrpIleGlyAsnSerAspLeuSerProGl	
1	GGACATCAGCCGCGGCAGTTCGTCATCAACCTGATGACCGAAGCCATGGCTCCGACCAACACCCTGTCCAACCGGCAGCAGTCAAACGCTTCTTCGAA nAsplleSerArgGlyGlnPheVallleAsmLeuMetThrGluAlaMetAlaProThrAsnThrLeuSerAsnProAlaAlaValLysArgPhePheGlu	
1	MCCGGGGGCAAGAGCCTGCTCGATGGCCTGTGCAACCTGGCCAAGGACCTGGTCAACAACGGTGGCATGCCCAGGCGAGGTGAACATGGACGCCTTCGAGG ThrGlyGlyLysSerLeuLeuAspGlyLeuSerAsnLeuAlaLysAspLeuValAsnAsnGlyGlyMetProSerGlnValAsnMetAspAlaPheGluV	
•	TOGGCAAGAACCTGCCCACCACCTGA BCCCCCACCTGA BCCCCCACCACCACCACCACCACCACCACCACCACCACCA	
_	CCCCTGCTGGTGGTGCCGCCAGATCAACAAGTTCTACGTATTCGACCTGAGCCCGGAAAAGAGCCTGGCACGCTACTGCCTGC	
7	CCTTCATCATCAGCTGGGGCAACCGGACCAAAGCCCAGCGGGAATGGGGCCTGTCCACCTACATCGACGGGGGGGG	
1	CATTACCGGCAGCAAGGACCTGAACATGCTCGGTGCCTGCTCCGGCGGCATCACCTGCACGGCCATTGGTCGGCCACTATGCCGCCCTCGGCGAAAACAA AlleThrGlySerLysAspLeuAsnMetLeuGlyAlaCysSerGlyGlyIleThrCysThrAlaLeuValGlyHisTyrAlaAlaLeuGlyGluAsnLy	
- G	GTCMATCCCCTGACCT TCC TTCCTCCACCCTCTCCACCCCCCCCCC	
_		
_	ACTECTACEAGGEGGGGGGGGAGGGGAGGGCGAGGATGGCCAAGGTGTTCGCCTGGATGCGCCCCAACGACCTGATCTGGAACTACTGGGTCAACAACT isserTyrGlnAlaGlyValLeuGluGlySerGluMetAlaLysValPheAlaTrpMetArgProAsnAspLeuIleTrpAsnTyrTrpValAsnAsnT	
λ. Υ.	CCTGCTGGCCAAGGGGGGGGGGGTTCGACATCCTGTTCTGGAACAAGGACACCACCGCCTGCGGGCGG	
_	MGAGCAACCCCTGACCCGGCACGCCCTGGAGGTTTGCGGCACTCCGATCGACCTGAAACAGGTCAAATCCGACATCTACAGCCTTGCCGGCACC LysSerAsnProLeuThrArgProAspAlaLeuGluValCysGlyThrProIleAspLeuLysGlnValLysCysAspIleTyrSerLeuAlaGlyThr	•
	ACGACCACATCACCCCGTGGCAGTCATGCTACCGCTGGCGCACCTGTTCGGCGGCAAGATCGAGTTCGTGCTGTCCAACAGCGGCCACATCCAGAGCA STASPHislleThrProTipGlnSerCysTyrArgSerAlaHisLeuPheGlyGlyLysIleGluPheValLeuSerAsnSerGlyHislleGlnSerI	:
	CTCAACCCGCCAGGCAACCCCAAGGCGGCGCTTCATGACCGGTGCCGATCGCCCGGGGGACCCGGTGGCCTGGCAGGAAAACGCCCACCAAGCATGCCGA LeuAsnProProGlyAsnProLysAlaArgPheMetThrGlyAlaAspArgProGlyAspProValAlaTrpGlnGluAsnAlaThrLysHisAlaAs	2
C1 pS	CCTGGTGGCTGCACTGGCAAAGCTGGCTGGGCGAGCGTGCCGGGGAGCTGAAAAAGGCGCGACCCGCCTGGGCAACCGTGCCTATGCCGCTGGGGAG EXTTPTTPLeuBisTxpGlnSerTrpLeuGlyGluArgAlaGlyGluLeuLysLysAlaProThrArgLeuGlyAsnArgAlaTyrAlaAlaGlyGlu	2
GC A.)	ATCCCCGGGCACCTACGTTCACGAGCGTTGAGCTGCAGCGCCGTGGCCACCTGGGGGACGCCACGGTGTTCATTTCACCCCATGAGTCACGCCCATGC ASerProGlyThrTyrValHisgluarg MetP	2

## FIGURE 6 (page 2)

2401	COCAACCCTACATCTTCAGGACCGTGGACCAGCCAGGTCCATCCGCACCCGGGCCCAAACCGCACCTGACCCGGTGCTGATCTTCAA  ROGInProTyrIlePheArgThrValGluLeuAspAsnGlnSerIleArgThrAleValArgProGlyLysProHisLeuThrProLeuLeuIlePheAs	24
2402	COGCATCOGTGCCAACCTGGACCTGGTGTTTCCGTTCATCGACCCACTGGACCCGGACCTGGACTCATTGCCTTTGACGTACCCGGGGTCGCCGCTCCACGCTCGCACCTGGACCTACCCGGGTCGCCGCCTCCACGCTACCCGGGTCGCCGCTCCACGCTACCCGGGTCGCCGCTCCACGCTACCCGGGTCGCCGCTCCACGCTACCCGGGTCGCCGCTCCACGCTACCCGGACCTACACACGACCTACACGACCTACACGACCACACGACCTACACACGACCTACACGACCACACGACCTACACACGACCTACACACAC	25
<b>2</b> 501	TORONOMIC	
2601		
2701	CTTGGGGCGCGCCCTGGCCCAGCAGTTCGCCCACGATTACCCCGAACGCTGCAAGAAACTGGTGCTGGCCGCCACCCGCAGCCGGTGATGGTGCCGGTTTpGlyGlyAlaLeuAlaGlnGlnPheAlaHisAspTyrProGluArgCysLysLysLeuValLeuAlaAlaThrAlaAlaGlyAlaValPetValPr	27
- 101	AGGEAAGCCCAAGGIGIIGIGGATGAIGGCCAGGCCCACGGCGIIACGIGCAGCCGATGAGCATGATCCCCCATTGCCCCCGACGATCTAIGGCGGCGGCTTCCGCGCGATGATGATGCGGCGCGCTTCCGCGATGATGATGATGCGGCGGCGCTTCCGCGATGATGATGATGATGATGATGATGATGATGATGATGATG	28
2601	COCCOTCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCOCCHOCCHOCCHOCOCCHOCCHOCCHOCCHOCCHOCCHOCCACCHOCCHO	290
2901	GCATCCATTCCATCACAACAACAACAACAACAACAACAACA	30/
3001	GCATCCACTGCTGCACAAGATCCAGCACCCTGGTGCTGGCCGGCGACGACGACGACGGTGATCAACATGCGCCTGCTGGCCTGGCGGGGGTLAGGATCAACATGCGCCTGCCTGGCCTG	300
1101	GATTCCCAATGCCCAGCTACACATTATCGACGACGGTCATTTGTTCCTGATCACCGGGCCGGAGGCCGTCGCCCCGATCATCATGAAGTTCCTTCAGCAA glleProAsnAlaGlnLeuHisileIleAspAspGlyHisLeuPheLeuIleThrArgAlaGluAlaValAlaProIleIleMcLysPheLeuGlnGln  CAA I	310
201	GAACGACAGCGCCCCTCATGCACCCTCGCCCGGCTTCGGGCGGG	320
301	ACCIGIGGGAGGGGGCATGCCCGGGAAGGTCTCGACAGCGAAATGGCTTAGACGAGGGAGTGTTGCCATGAAAGGACAAACCGGCCAAAGGAACGCCCAACG HetProAlaLysValSerThrAlaLysTrpLeuArgArgGlySerValAlaMetLysAspLysProAlaLysGlyThrProThr	330
	CTTCCCGCCACCAGCATGAACGTGCAGAACGCCATCCTCGGCCTGGCGGGGGGGG	340
501	ACCCGCTGCACACCGCACATCACCTGTTGGCCCTGGGTGGCCACCTGGGCCGGTGATACTGGGTGACACACCGCTTCAGCCGAACCCGCATCCGG isProLeuHisThrAlaHisHisLeuLeuAlaLeuGlyGlyGlyGlnLeuGlyArgValIleLeuGlyAspThrProLeuGlnProAsnProArgAspProAr	350
501	CITCAGCGACCCGACATGGAGCCAGAACCCGTTCTACCGGCGCGCCCTGCAAGCCTACCTGGCCTGGCAGAAGCAGACCCGCCTGTGGATCGAGGAAAGC gPheSerAspProThrTrpSerGlnAsnProPheTyrArgArgGlyLeuGlnAlaTyrLeuAlaTrpGlnLysGlnThrArgLeuTrpfleGluGluSer	360
	CACCTOGACCACGATGACCGGGCCCGGTGCGCACTTCCTGTTCAACCTGATCAACGATGCCCTGGCGCCAAGCAACTCGCTGCTCAACCCGGTGGCGGTCA Hislewaspaspaspaspargalaargalahispheleupheasnleulieasnaspalaleualaproserasnserleuleuasnproleualavall	370:
	Aggaactettcaacagggggggggggggggggggggggggg	3800
	GCGCGCCTTCGAAGTGGGCGGCAACCTGGCCGCGACTGCCGGCGCCCCGTGGTGTTTCGCAACGAGCTGCTGGAACTGATCCAGTACAAGCCGATGAGCGAA WArgAlaPheGluValGlyGlyAsnleuAlaAlaThrAlaGlyAlaValValPheArgAsnGluLeuGluLeuGluLeuTleGlnTyrLysPromecSerGlu	3900
٠- ١	MAGCAGCACGCCCGGCCACTGCTGGTGGTGCCGCCACAGATCAACAAGTTCTACATCTTCGACCTCAGCTCGACCAACAGCTTCGTCCAGTACATGCTCA LysGlnHisAlaArgProLeuLeuValValProProGlnIleAsnLysPheTyrIlePheAspLeuSerSerThrAsnSerPheValGlnTyrHetLeuL	4000
- 01 ر	AGRATICA TO THE TOTAL TO	<b>4</b> 100
01 -	ACCOTTOTOTACENTESCOCON	
1 -	AACGCTTGCCGCAGCATTAGCGGCAACCGCGACCCCAACCTGATGGCCGCCTGCGCGGGGCCTGACCATGGCCGCACTGCAGGGCCACCTGCAGGCCACCTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAGGCCAACTGCAACTGAACTGAACTAACT	4200
1 -	ysHisGlnLeuArgArgValArgSerAlaThrTyrLeuValSerLauLeuAspSerLysPheGluSerProAlaSerLauPheAlaAspGluGlnThrI	4300
1 -	CGAGGCCGCCAAGCGCCGCTCCTACCAGCGGGGTGTGCTGGATGGCGCCGAGGTGCCCGGGATCTTCGCCTGGATGCGGCCCCAACGACCTGATCTGGAA eGluAlaAlaLysArgArgSerTyrGlnArgGlyValLeuAspGlyAlaGluValAlaArgIlePheAlaTrpHetArgProAsnAspLeuIleTrpAs	4400
C n	TACTGGGTCAACAACTACCTGGCTCGGCAAGACACCAGCCTTCGACATCCTGTACTGGAACGCCGACAGCACGCGCCTGCCCGGCGGCTGCATGGC TyrTrpValAsnAsnTyrLeuLeuGlyLysThrProProAlaPheAspIleLeuTyrTrpAsnAlaAspSerThrArgLeuProAlaAlaLeuHisGly	4500
1 -		4600

## FIGURE 6 (page 3)

ŧ

	GACCTOCTGGACTTCTTCAAGCTCAACCCGCTGACCCACCCAGCCGGGGGGGG
4601	•
	TCACCGTGGCCGCACCACGACCACATCACCCCGGGGGATCCGGTGTACCGCTGGGCCTTGCTGCTGGGGGGACCGGGGCTTCGTGCTGGCCAACAG heThrValAlaGlySerAsnAspHisIleThrProTrpAspAlaValTyrArgSerAlaLeuLeuGlyGlyAspArgArgPheValLeuAlaAsnSe
4701	
	CGGCCACATCCAGAGCATCATCAACCCGGCCAACCCCCAAGGCCTACTACCTGGCCCAACCCCAAGCTGTCCACCGACCCGGCTGCCTGC
4801	
	CCCANCCCACCCAACCCCACCCCACCCCACCCCACCCC
4901	
	CCTACCCACCGCCCCCCCCCCCCCCCCCCCCCCCCCCCC
5001	
E101	AGCTGTTCAACCAGCAGGGGGAACCGAACGTATCCACCCTGGAAATTGCCAACGGAACTGGGCATCAGCCCTGGCAACCTCTACTACCACTTCCACGGCAA
5101	GGAGCCGTTGGTGCTGCTGCTGCTTGTTGACCACCGCTTTTTAACAACCGCTTGTTGACCACCGCTTGTTGACCACCGCTTGTTGACCACCGCTTGTTGACCACCGCTTGTTGACCACCGCTTGTTGACCACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACAACCGCTTGTTGAACCAACC
5201	
	CTGTTCCTGCACCTGATCGTCGAACGCATGGCGCAGTACCGCTTCCTGTTCCAGGACCTGTCGAACCTGACCGGCGCCCTGCCCAAACTGGCCCGCCAC
5301	TOTALTICATICATICATICATICATICATICATICATICATICA
5401	TGCGCAACCTGATCAACGCGCTCAAGCGCACACTGGCGGGGTTGCTGGCCAGGCCTCAAGGGCTTGGTAGAGAGCGGAGACCCAGGCGCTGGGGCA
	ACTOGTGGAGCAGATCACCCTGACACTGATGTTCTCGCTGGATTATCAGCGGGTACTGGGGGCATGTGGGGGATTGTGGTGTACCAGTGATG
5501	
5601	ATTICTOGRACOCCATCTICAGGCCCAGGGCCGGGGGGGGGGGGGGGGGGGG
	CTTTCCCCGGCATGCCCCCACAGGTGAAATGCAGTGCTCGAGTGCACACAGGACCTGTGGGACCGGGCAAGCCCCGGAAGATGGCGGACGCGGTATC
701	
	AGATCAGGGTACCGGTGCCTGTCTGTGCCGAAGGCGGGTTGCTGGCCGGAGGCCGCAGTGGCCGCGAAGCTGCCCCTAGCCGCCGGAAGCTGGCCGC
801	TGGACCTCCCCCTTCCCCCTTCCCCCTTCCCCTTCCCCTTCCCC
901	TOGACCIGCOGCTIGGCCGCCTTGCCCGGTGCTTTCTTCACTGCAGGCTTCTTCGCCACAGGCCGTTTGGCCCGCTTTCGCCCCCAGGC
	TIGGCCGCCAGICTIGGCAGCAGGTTTAGCCGCTIGGCGGCCTTGGCCGCGGTTTTTGCCGCAGGCTTGGCCGGTCTTGG
5001	CCAGTOCCTTCCTTCCTTCCTCCTCCTCCTCCTCCTCCTCCTCC
101	CCAGTGGCTTGGCCGCCGCCTTGCTCGCAGCCGGTTTGGCTGCAGTGCGGCACGAAATCGGGCGTAACCGAAGGGCCCGTGAGTTTCTCGATCTGGTTGGT
	CAGGCTGTCCAGCTGCTGGTGCAGGGCCTTGATCTGGTTGCGGCTGGGCACGCCAAGGCGCGAGATGGCACTGTTCAGGCGCTTTGTCGAAGGCCTCTTCG
201	
301	AGTTCCCTCCACTTCCCCTACCCCCCCCCCCCCCCCACACACCCCCAAAGTCGTCCACGACTTCGCCAGTTTCAGCAAACATCTTCTCCCGGTCTTCTTC
	GCCTGTTTCTGGGCTTCTGGTATTCTTTCTACTACTACTACTACTACTACTACTACTACT
401	GOCTETTTCTCGCCTTCTCGCCATCCTTTACCAGCGAGTCGAACAGCTTCGGGCCGTCCTGGTCGATCTTCGAATAGATACCAAGCCCCCCCC
	TCTTGCGGGAGTACTTCTCGATCCCGCCGACCCAGGAGCTGCCTTCTTCTCGGAATTC

## INTERNATIONAL SEARCH REPORT

International Application No PCT/US 90/03851

1. CLASSIFICATION OF SUBJECT MATTER (it several classification symbols apply, indicate all) 6								
According to International Patent Classification (IPC) or to both National Classification and IPC								
IPC <sup>5</sup> : C 12 P 7/62, C 12 N 15/52								
II. FIELDS	SEARCHED							
	Minimum Document	ation Searched 7						
Classification System   Classification Symbols								
IPC <sup>5</sup>	C 12 N, C 12 P		-					
Documentation Searched other than Minimum Documentation to the Extent that such Documents are included in the Fields Searched •								
IIL DOCUMENTS CONSIDERED TO BE RELEVANT								
Category •	Citation of Cocument, 11 with Indication, where appr	opriate, of the relevant passages 12	Relevant to Claim No. 13					
x	WO, A, 89/00202 (MIT) 12 January 1989 see the whole docum		1-8,10,11, 25-27					
P,Y	(cited in the applicati  The Journal of Biologic		•					
•	volume 264, no. 26, The American Societ and Molecular Biolo O.P. Peoples et al. butyrate Biosynthes eutrophus H16", pag see the whole artic	1-8,10,11, 25-27						
P,X	The Journal of Biologic volume 264, no. 26, The American Societ and Molecular Biolo O.P. Peoples et al. hydroxybutyrate (PH Alcaligenes eutroph	1-8,10-15, 20-22,25-27						
"A" document defining the general state of the art which is not considered to be of particular relevances "E" sariier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to setablish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but leter than the priority date claimed  IV. CERTIFICATION  Date of the Actual Completion of the International Search  27 NOV 1990								
Internation	nai Searching Authority	Signature of Authorized Officer						
	EUROPEAN PATENT OFFICE	Mm Mm	N KI IIDED					

Calegory *	CHARGO OF DOCUMENT, 11 with Indication, where appropriate, of the relevant passages	Relevant to Claim No.
Tradory -	Custion of Document with indication' where appropriate' or the season's besteven	Relevant to Claus No.
	pages 15298-15303	
}	see the whole article	
[	1	]
1	en e	
X	Journal of Bacteriology, volume 170,	1,2,7,8,10,
ł	no. 12, December 1988, American	11,25-27
	Society for Microbiology,	
	P. Schubert et al.: "Cloning of the	ļ
	Alcaligenes eutrophus genes for synthesis of poly-6-hydroxybutyric	
	acid (PHB) and synthesis of PHB in	
ļ	Escherichia coli", pages 5837-5847	
	see the whole article	
. 1		
x	Journal of Bacteriology, volume 170,	1,2,7,8,11,
	no. 10, October 1988, American	25-27
ŀ	Society for Microbiology, S.C. Slater et al.: "Cloning and	
	expression in Escherichia coli of the	
İ	Alcaligenes eutrophus H16 poly-P-	İ
	hydroxybutyrate biosynthetic pathway",	
	pages 4431-4436	
	see the whole article	
•	<b>Cas Cas</b>	
T	Science, volume 24, 15 September 1989,	
	(Washington, DC, US),	
	R. Pool: "In search of the plastic	
	potato", pages 1187-1189	
	see the whole article	•
	· •••	
A	Chemical Abstracts, volume 110, no. 15,	
	April 1989, (Columbus, Ohio, US),	
	G.W. Haywood et al.: "The importance	
	of PHB-synthese substrate specificity	
	in polyhydroxylalkaneate synthesis	,
	by Alcaligenes eutrophus",	
	see page 404, abstract 13192p & FEMS Microbiol. Lett. 1989, 57(1),	
	1-6 (Eng).	
	(	· ·
		ľ
-		
	•	
	7. No. 1	

### ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO.

US 9003851 SA 38887

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 13/11/90

The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Pater men	nt family nber(s)	Publication date
WO-A- 8900202	12-01-89	EP-A-	0329770	30-08-89
	•			•
	•			•
		•		
•	·			
			•	
•				
	:			
			•	
		•		